

OTDK-dolgozat

2025.

Budapesti Corvinus Egyetem

**Számít-e a monetáris politika? A magyar monetáris sokkok azonosítása és
empirikus elemzése szövegelemzéssel**

**Does Monetary Policy Matter? Identification and Empirical Analysis of the
Hungarian Monetary Shocks via Text Analysis**

Kovács Márton Attila

Konzulens: Reguly Ágoston

Kézirat lezárásának dátuma: 2024. 11. 21.

I Introduction

Monetary policy actions, taken by the central banks, are key to achieving price stability, through the control of money supply and nominal interest rates. These actions generate monetary shocks, which are hard to determine through conventional methods and are aimed at influencing economic indicators in order to push the economy back to the equilibrium path. This analysis aims to provide a framework that identifies these shocks via state-of-the-art machine learning techniques, through which the impact of monetary shocks is analyzed on the real economy. As a response to contractionary monetary shocks, conventional methods signal a 0.4 percentage point rise in unemployment and a 0.5% drop in inflation, while using the measures developed in this analysis, unemployment may rise by more than 1.5 percentage points with inflation declining by 2%.

Previously the identification of monetary shocks by text analysis has been achieved by analyzing publications from Central Banks, as shown by Romer and Romer (2023). Although this technique is reliable and precise, it requires a significant amount of time and effort from various research groups. Another more frequently used technique is identifying monetary shocks based on asset prices. This conventional method may be subject to bias owing to unquantifiable market dynamics. The objective of this study is to address identification challenges by using machine learning and text analysis techniques. This approach, while demanding significant computational resources, aims to minimize tedious human tasks, eliminate biases, and help researchers focus on interpretation and fine-tuning. Furthermore, the analysis framework and methods can easily be generalized to any monetary text, with the limitation that the text should be written in English, thus allowing the analysis to be conducted for any country where monetary publications are released in English.

To conduct the analysis, the literature was reviewed in three major areas. First, research on text analysis in a financial context has been examined, followed by central bank communication and shock analysis.

The development of Bidirectional Encoder Representations from Transformers (BERT) models has lately allowed large-scale language modeling. These models have far surpassed earlier models in their capacities for general-purpose applications (Devlin et al., 2018). In less than a year the BERT model was fine-tuned on financial datasets, for better accuracy when working with financial

data (Araci, 2019). Simultaneously, the RoBERTa model was introduced, which was a better-optimized version of the original model (Liu et al., 2019). The introduction of BERT models has facilitated the development of few-shot learning algorithms, which use these models to perform text classification tasks with a small amount of labeled data, one such model is SetFit (Tunstall et al., 2022). Chat GPT was introduced with the advancement of LLMs (Large Language Models), showcasing impressive zero-shot learning skills. This progress continued with the subsequent versions of Chat GPT, namely 3.5 and 4 (Brown et al., 2020, OpenAI, 2023, Koubaa, 2023 and Ye et al., 2023). Recently, several research publications have examined the performance of these models in financial tasks. In their study, Shah and Chava (2023) conducted a comprehensive evaluation of zero-shot learning models across several financial tasks. Their findings indicate that, while Chat GPT performs better than other zero-shot learning models, pre-trained fine-tuned models like RoBERTa surpass it. In addition, Loukas et al. (2023) discovered that both Chat GPT and SetFit models provide favorable classification outcomes when evaluated using a financial dataset. Shah et al. (2023) conducted a comparative analysis of several language models for classifying phrases extracted from publications by the FED. The researchers discovered that a RoBERTa model, which has been fine-tuned via the process of manually labeling texts, produces the most favorable outcomes. However, both Chat GPT and simple rule-based models also show encouraging results. Additionally, a simple method that relies on identifying terms within the data was examined, using a financial dictionary developed by Gorodnichenko et al. (2021).

Furthermore, papers on monetary communication were examined, with the theory behind it. Blinder et al. (2008) provided a survey of theory and evidence on central bank communication up to that point. The basic theory of communication was laid out, with an empirical analysis of the effect of communication on financial and economic variables. Since then many good research papers have been released focusing on the ECB (De Haan and Sturm, 2019, Blinder et al., 2022), FED (Romer and Romer, 2023) and emerging markets including Hungary (Evdokimova et al., 2023, Rozkrut et al., 2007 and Gábríel and Pintér, 2006). Research based on large-language models in the context of central bank communication, to our best knowledge, has only been conducted by Shah et al. (2023), as discussed previously. After reviewing the papers, two general conclusions can be made, i) transparent central bank communication is key, and ii) communication has an effect, and can be used to proxy shocks.

Lastly, a short overview of shock modeling is given. After the release of the seminal paper “Macroeconomics and reality” by C.A. Sims, and the introduction of Vector Autoregressive models (VARs), economics were able to calculate the impulse-response functions of macroeconomic variables to shocks (Sims, 1980). These functions showed the impact of shocks on an economy and helped policy actions. Later, Jordá proposed a system, where the impulse-response functions could be calculated, without computing the underlying multivariate system (Jordá, 2005). This enables the calculation of the response function using a series of OLS regressions. Romer and Romer (2023) have used the method proposed by Jordá to calculate the effects of monetary shocks on the economy. They identified shocks by reading minutes publications by the FED.

In this study, several methodologies that were previously described are used, both in textual analysis and shock analysis. The paper presents a zero-shot learning model that utilizes Chat GPT 3.5, a few-shot learning approach based on SetFit, a fine-tuned model, and a rule-based methodology (Shah et al., 2023, Shah and Chava, 2023 and Gorodnichenko et al., 2021). Local projections are used for shock modeling following Jordá (2005), a modeling framework utilizing a sequence of linear regressions to produce estimates. Using these techniques, four sentiment measures are computed for the Hungarian Central Bank. By utilizing these metrics as indicators of shocks, supplemented with a fifth measure derived from foreign exchange rates, the study examines the response functions of real-economic variables. The analysis concludes that monetary shocks have significant impacts on the Hungarian economy, leading to higher unemployment and lower output. Nevertheless, the monetary policy proves to be efficacious in addressing inflation after the transmission period, resulting in a reduction of up to 2 percentage points.

The paper is organized as follows. Section II presents the data used in the study, followed by a comprehensive explanation of shock identification techniques in Section III. In Section IV the results of the shock analysis are discussed. Section V concludes.

II Data

For the study, two types of data are used regarding the Hungarian economy. First, real economic and financial data are presented, and as a second step, we introduce the monetary-related text documents.

II.1 Real-economic and financial data

For the analysis, data was collected from both the real economy and financial sector, covering three different frequencies: quarterly, monthly, and intra-day. Data on real GDP and inflation were collected from the Hungarian National Bank's dataset¹, which is used for their inflation reports. The real GDP is measured in Hungarian Forints and is based on constant 2005 prices. Both the Consumer Price Index (CPI) and core inflation are commonly utilized to gauge inflation. Using both inflation measures is important because core inflation excludes the most volatile items, which can help mitigate the impact of a monetary shock (KSH, 2000). When analyzing monthly data, we consider the average euro-forint exchange rate for each month and the nominal base rate set by the central bank. The data is sourced from the MNB's statistical databases (MNB, 2024). In addition, seasonally adjusted unemployment statistics are collected from Eurostat as a percentage of the labor force (Eurostat, 2024). Finally, a Euro-Forint exchange rate was collected at half-hour intervals from Bloomberg. The opening rate is used for all intervals. The data has been collected from 2008 until the end of 2023, with the exception of real GDP and inflation, which were last measured in 2023 Q2.

II.2 Text data

To analyze monetary policy sentiment, we opted to utilize the Minutes of the Monetary Council Meetings (MMCM). These documents are released every two weeks following each rate-setting Monetary Council Meeting. While there is a short explanation released with the rate decision after every rate-setting meeting at 2 pm, the MMCMs provide a more comprehensive stance of the Monetary Council. The extended documents include the stance of monetary policy on a wider range of issues such as macroeconomic developments, and adds the possible counterarguments to the rate decision by board members in the minority, which is not available in the short explanation. In total 224 documents were downloaded from the MNB's website ranging from 2004 December to 2023 December. There were twelve documents per year and all documents were collected in English.

Once the data was downloaded, the documents were converted into a plain text format, and

¹<https://www.mnb.hu/kiadvanyok/jelentesek/inflacios-jelentes/2024-03-28-inflacios-jelentes-2024-marcius>

any recurring sections that appeared at the beginning of each document were removed. After identifying the main part of each document, the sentences were split and written into separate Excel files. A total of 18,695 sentences were captured, resulting in 548,216 words, over 19 years. On average, each main body of a document contained about 83 sentences or approximately 2,500 words, resulting in an average of 29 words per sentence. In addition, the documents had an average Flesch-Kincaid grade level score of 14.7, suggesting a level of readability typically found in university-level texts. While the number of words and sentences across documents stayed relatively the same over time, the Flesch-Kincaid grade level score varied over time.

Variable	Description	Years Available	Source
rGDP	Real GDP measured in 2005 constant domestic prices	2008-2023Q2	MNB inflation report database
CPI	Consumer Price Index	2008-2023Q2	MNB inflation report database
CCPI	Core inflation	2008-2023Q2	MNB inflation report database
Unemployment	Unemployment rate as a percentage of the labor force	2008-2023	KSH
Base rate	Nominal interest rate set by the MNB	2004-2023	MNB statistical database
EURHUF	Euro-Forint exchange rate at monthly and intra-day frequency	2008-2023	MNB statistical database and Bloomberg
MMCM	Minutes of the Monetary Council Meetings	2004-2023	MNB

Table 1: Data used during the analysis

Notes: real GDP, CPI, and CCPI are quarterly series, while unemployment, base rate, EURHUF, and MMCM are monthly.

III Shock identification

This section provides a description of the two unique methods utilized in the study to identify shocks. First, a classic shock identification based on the intra-day EURHUF exchange rate, which reacts heavily to monetary policy in an inflation-targeting regime (see, Galí and Monacelli, 2005

or Bjørnland, 2008), and then the natural language processing techniques are discussed (Tunstall et al., 2022, Shah et al., 2023, Shah and Chava, 2023 and Gorodnichenko et al., 2021).

III.1 Classical identification

As mentioned, conventional methods for identifying shocks rely on using the EURHUF exchange rate, which is known for its significant daily trading volume. Moreover, the euro-forint exchange rate serves as a reflection of investors' perception of the country and its economy. Therefore, it is highly probable that a monetary shock will have an immediate impact on this market, given the fast-paced trading capabilities. The data was analyzed by focusing on the rate-setting meetings of the Monetary Council. It was then divided into two groups: the rates before the meetings and the rates after the meetings. When examining the exchange rate ER_t for potential structural breaks, we utilized a step-dummy variable D_t that had zeros prior to the rate announcement. This variable was then regressed on the combined exchange rate at time t , defined by Equation (1),

$$ER_t = \beta_{0,t} + \beta_{1,t}D_t + e_t. \quad (1)$$

When analyzing monetary shock in this way, one identifies a monetary shock at time t if $\beta_{1,t}$ is significantly different from 0. In addition, the sign of $\beta_{1,t}$ indicates whether the shock had a contractionary effect (strengthening the exchange rate, thus a negative sign) or an expansionary effect (weakening the exchange rate, thus a positive sign), supported by Bjørnland (2008). If a shock is identified as contractionary, a value of 1 is assigned to it, while if it is identified as expansionary -1 is assigned. All dates without identified shock are assigned the value 0. As a further measure, we calculate linear trends for the two groups of exchange rates to evaluate any changes in *trends*. Following the calculations, the euro-forint exchange rate on all rate-setting meeting days was plotted and examined. If there is a structural break in the exchange rate at time t , followed by a trend change immediately after the announcement at 2 pm, it would be considered a monetary shock.

Following this framework for identification, a total of 32 shocks were identified, with 14 contractionary and 18 expansionary. Most of the shocks occurred in the years prior to 2010 and after 2020. However, there was at least one shock every single year, except for 2017. Figure 1 shows

a contractionary monetary shock on 2007 July 21st. There was a significant trend change around the announcement of the interest rate decision and publication of a short explanation, while the coefficient of the step-dummy regression was significantly different from 0.

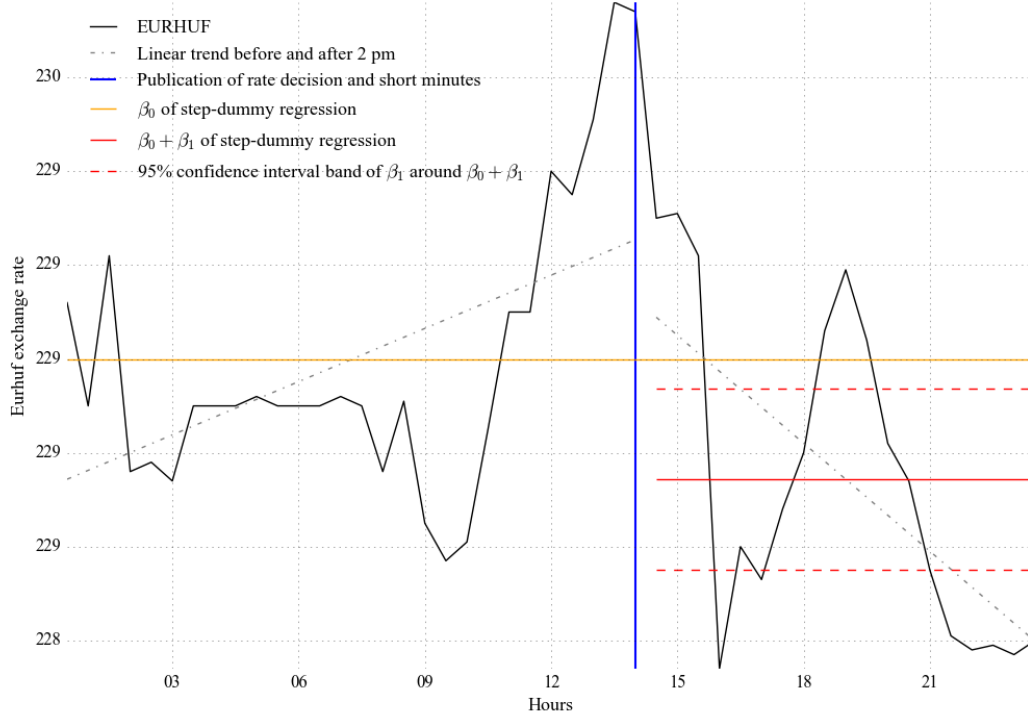


Figure 1: Contractionary monetary shock on 2007 July 21st

Notes: A contractionary monetary shock on 2007 July 21, resulting in the strengthening of the Hungarian Forint. Alongside the monetary announcement at 2 pm, two linear trends are plotted against the exchange rate before and after the announcement. The two coefficients of the step-dummy regression are plotted, and the 95% confidence interval band is shown for β_1 around $\beta_0 + \beta_1$.

Although this technique is straightforward in terms of computational complexity and comprehensibility, it does not account for the intensity of disturbances and fails to recognize disturbances when other market factors may have caused the market to move in the other direction, so nullifying the impact of the disturbance. To address these concerns in the next part, a novel method, based on text analysis, of quantifying shocks is introduced.

III.2 Identification with language processing

Although classical shock identification has its merits, it fails to fully capture the magnitude of the shock’s impact on the real economy. Here we employ several natural language processing techniques to assess the MNB’s communication and generate a continuous variable that captures the magnitude of shocks. The models classify monetary *sentences* into contractionary, expansionary, or neutral, and then create a “score variable” for each *document*, which reflects the monetary stance. Equation (2) describes how we assign the score for each document i based on each sentence s in document i .

$$Score_i = \frac{\#Contractionary_{i,s} + \#Expansionary_{i,s}}{\#Total_{i,s}}. \quad (2)$$

To classify each sentence, four different techniques were employed: a rule-based approach, zero- and few-shot learning, and a fine-tuned RoBERTa model. In this section, each of the four models are discussed.

First, a rule-based approach was adopted, which has been applied in other research previously, and is specifically designed for monetary texts (Gorodnichenko et al., 2021, Shah et al., 2023). We used a total of five sets of words, and an algorithm was provided with guidelines for assigning labels depending on the presence of terms in each sentence. See the guidelines for classification in Table A.1 in the Appendix.

In addition, a zero-shot methodology, where a language model was not explicitly trained on any monetary data, was implemented using ChatGPT 3.5, a language model that has been utilized in several recent academic studies exploring financial communication and its zero-shot capabilities (Loukas et al. 2023, Shah et al. 2023 or Shah and Chava 2023). The models were presented with the following prompt, for each sentence and were asked to classify the sentence: “Discard all the previous instructions. Behave like you are an expert sentence classifier. Classify the following sentence from the Hungarian National Bank into ‘HAWKISH’, ‘DOVISH’, or ‘NEUTRAL’ class. Label ‘HAWKISH’ if it corresponds to the tightening of the monetary policy, ‘DOVISH’ if it corresponds to the easing of the monetary policy, or ‘NEUTRAL’ if the stance is neutral. Provide the label in the first line and provide a short explanation in the second line. The sentence: {sentence}” (Shah et al., 2023).

Thirdly, we used a few-shot learning approach, where a model can generalize only after training

on a few examples per class, called SetFit described in (Tunstall et al., 2022). This method utilizes few-shot fine-tuning of pre-trained large language models, using a contrastive Siamese method. This method can achieve results that rival fine-tuned models, with only 8 labeled sentences per class. Loukas et al. (2023) found promising results using this framework on financial classification tasks.

Finally, a fine-tuned, pre-trained language model is used, called RoBERTa, which was first trained on large amounts of text data and was later fine-tuned by Shah et al. (2023) on labeled text data by the Federal Reserve. While, this approach to sentence classification is the most sophisticated, a possible source of bias is in the discrepancy between the communication of the Federal Reserve (FED) and the Hungarian National Bank (MNB). Possibly, the most important difference is the double mandate of the FED, that its primary objective is price stability and full employment at the same time, while for the MNB its primary objective is price stability. This difference may influence the communication of the two central banks, especially, when the inflation objective is achieved.

Although this method produces a continuous variable that represents the position of the monetary governance, it assumes that the size of any disturbance will be reflected in the size of the communication score. While this assumption is generally valid, it fails to account for instances when central banks exhibit dovish-tightening or hawkish-easing conduct, meaning their communication posture is contrary to their actual actions. This may result in noisy outcomes when a central bank is involved in such behavior.

III.3 Identified monetary shocks

The collected MMCM (Minutes of the Monetary Council Meeting) documents start from 2004 to 2023. Four distinct language models were run as described above. A negative score at time t means a majority of dovish sentences found in the MMCM document describing the meeting held at t , while a positive score means the opposite. While the sign of the score signals the overall sentiment, the magnitude of the score tells about the intensity of the sentiment. A score with an absolute value closer to one indicates a stronger intensity in the document, in the direction indicated by the sign of the score. The results are plotted in Figure 2, alongside the 5-month backward-looking moving

averages in Figure 3. While the majority of models have an average score close to 0, the zero-shot learning model, using Chat GPT 3.5, has an average score of -0.22, indicating on average a more dovish stance. Although all models have significant volatility, the rule-based and SetFit models demonstrate notably lower levels of volatility compared to the other two models. During the years before 2010 and beyond 2020, all models demonstrate an increasing sentiment trend, although the period in between is characterized by a more cautious and conciliatory stance. The rule-based and SetFit models have consistently yielded lower sentiment scores, in absolute values, throughout the years, indicating that they were unable to detect the shifts in monetary policy stance. As a result, they may not be as reliable in representing monetary shocks.

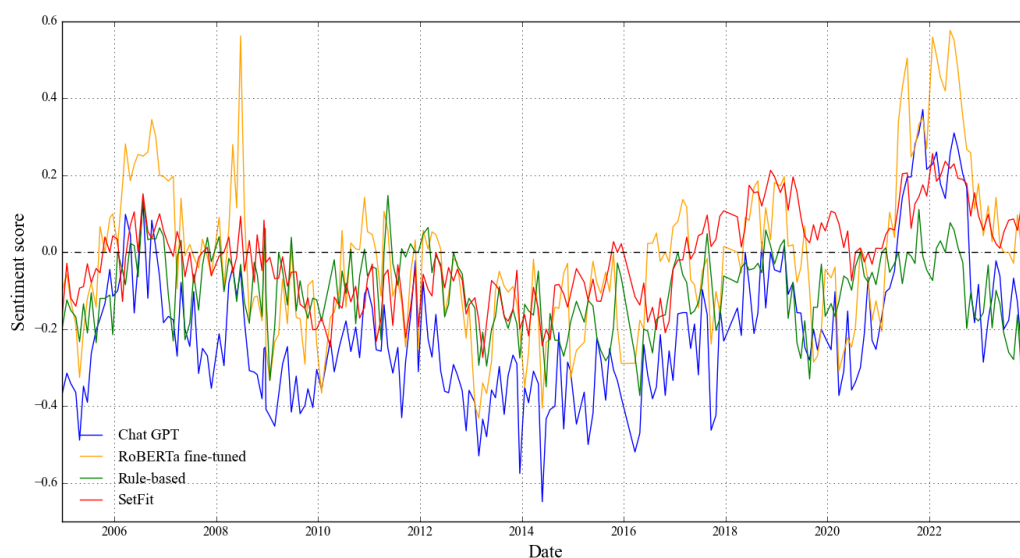


Figure 2: Sentiment scores of the Minutes of the Monetary Council Meeting documents

Notes: Sentiment scores of the MMCM documents from 2004 December to 2023 December. Higher scores refer to tightening sentiment, while lower score relates to easing of monetary policy.

To determine the optimal model, analyzing just the means and standard deviations is insufficient. Therefore, the models will be further examined based on their accuracy, correlation, and effectiveness as predictors. Accuracy is determined by comparing the categorized sentences with pre-labeled sentences. Furthermore, the analysis examines the link between the scores and macroeconomic and financial variables using Pearson correlation coefficients and Granger causality tests to determine their predictive capabilities.

In order to assess the accuracy of our results, six documents were manually labeled, at least

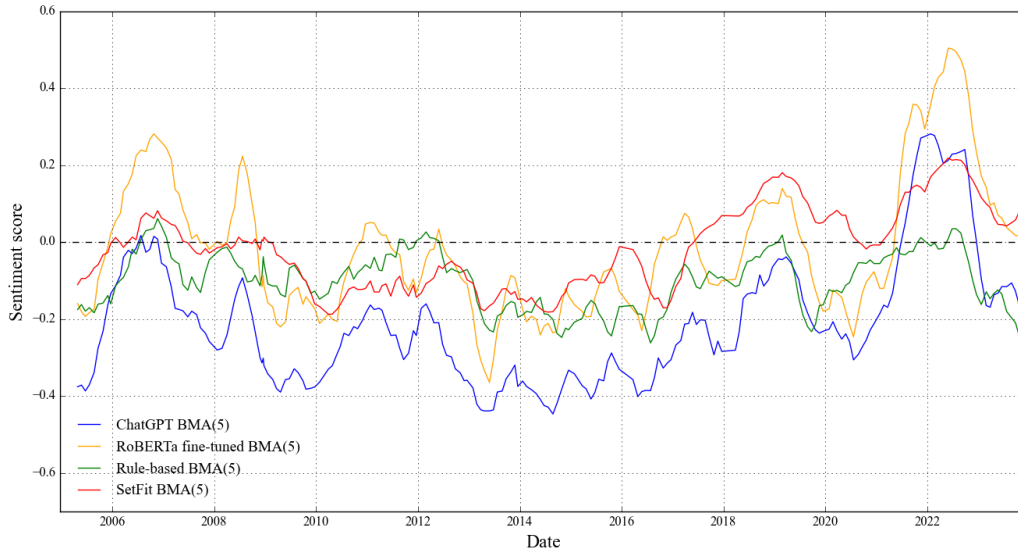


Figure 3: Smoothed sentiment scores of the Minutes of the Monetary Council Meeting documents
Notes: Backward looking five-month moving averages of sentiments scores of the MMCM documents from 2004 December to 2023 December. Higher scores refer to tightening sentiment, while lower score relates to easing of monetary policy.

one from each Central Bank Presidential, during the timeline of our analysis. After comparing the pre-labeled documents, with the language models and averaging out the accuracy scores across documents, we found a mediocre fit in terms of accuracy.²

Correlation coefficients were computed to assess the relationship between language model scores and financial and macroeconomic factors. The comprehensive collection of correlation coefficients may be seen in the Appendix. The models with the greatest correlation coefficient between them are Chat GPT and RoBERTa fine-tuned, with a value of 0.83, while the lowest is 0.31 between SetFit and rule-based. The rule-based also resulted in the lowest overall correlation with the other models for an average of 0.41, while the zero-shot model based on Chat GPT averaged a correlation coefficient of 0.68. As all four of these variables measure the same unobserved monetary shock, their correlation is expected to be significant, and the poor overall correlation of the rule-based model with the other measures means that it isn't able to capture the underlying

²The best accuracy was achieved by SetFit, with a mean accuracy of 40%. The second best model was the rule-based with 21%, closely followed by Chat GPT and RoBERTa fine-tuned with 18% and 17% respectively. While the accuracy scores may seem low, the choice of the six document and author-labeled sentences which may be biased can explain these low values. Note that a potential extension of this paper would be to label all the documents and do a more thorough analysis with further fine-tuning.

variable.

Another important check is to see how the created variable correlates with other economic/financial data. To show this, we calculated the correlation coefficients, which revealed a strong inverse relationship with unemployment, with an average coefficient of -0.38. However, this coefficient drops to -0.52 after eliminating the association with the rule-based model. The Euro-Forint exchange rate showed a connection of 0.33, whereas real GDP, CPI, and the change in the nominal interest rate set by the MNB had correlations of 0.59, 0.4, and 0.41, respectively. Despite the already elevated levels, when the rule-based estimations are excluded, the correlation coefficients increase to 0.45, 0.72, 0.52, and 0.45 correspondingly.

For the measure of the scores' effectiveness as forecasters we chose to use Granger causality tests, to determine whether macroeconomic or financial variables can be forecasted using our obtained values. The Granger causality test is a useful tool for assessing if a series, denoted as X_1 , has a Granger-causal influence on another series, denoted as X_2 . This means that the previous values of X_1 include valuable information that aids in predicting the future values of X_2 (Granger, 1969, Seth, 2007). In our case, this means that our measure has information relating to future values of macroeconomic and financial variables and, thus is effective in forecasting. Although this approach is not able to identify true causality, hence we use the term granger-causes, it is still an important measure to see if our calculated measure contains information on the chosen variables³. The null hypothesis posits that there is no causal relationship between the time series, while the alternative hypothesis states that one time series Granger causes the other. We have shown that all four series exhibit Granger causality with respect to the fluctuations in the nominal exchange rate determined by the central bank. This finding suggests that the sentiment of the Hungarian Central Bank at time t serves as a reliable signal for potential future changes in the exchange rate. Moreover, notable findings were observed regarding the Granger causation between the Chat GPT and RoBERTa models and the Euro-Forint exchange rate, unemployment, real GDP, and inflation. The SetFit model generated noteworthy findings for the majority of the variables, however, no Granger causality was observed for inflation. The rule-based model did not show statistical significance for any variable except for the change in the bank's interest rate.

³For further reading and the limitations on Granger causality the works by Stern (2011) and Shojaie and Fox (2022)

Variable	Chat GPT 3.5	RoBERTa fine-tuned	SetFit	Rule-based
Nominal Interest Rate Change	0.022 ** (3)	0.005 *** (2)	0.017 ** (3)	0.014 ** (2)
Euro-Forint exchange rate	0.015 ** (1)	0.051 ** (1)	0.019 ** (1)	No significant results
Unemployment	0.021 ** (1)	0.056 * (1)	0 *** (1)	No significant results
Real GDP	0.033 ** (2)	0.061 ** (2)	0.005 *** (1)	No significant results
CPI	0.03 ** (2)	0.056 * (2)	No significant results	No significant results

Table 2: Granger causality p-values

Notes: The table contains the first Granger causality p-value, which was significant at 10% or lower. The brackets show the corresponding lags. If no significant p-value is found in the first five lags, no value is shown. Nominal interest rate change, Euro-Forint exchange rate, and unemployment are monthly series, while real GDP and CPI are quarterly.

The statistical examination of the text analysis data revealed encouraging findings. Significant correlations and strong predictive capabilities were seen in the majority of models, with the exception of the rule-based model. These findings indicate that the traditional approach to text analysis may not provide satisfactory outcomes. The models have strong predictive capabilities for all variables and exhibit high correlation coefficients and their significant volatility may be beneficial for the modeling process. Upon reviewing the data, the RoBERTa fine-tuned and Chat GPT models were selected for the primary component of the shock study.

IV Real-economy effects of monetary shocks

In this section, we utilize the monetary sentiment and classic shock identification variables to compute the impulse-response functions of macroeconomic variables, employing the Jordá local projection approach. The response of the variables will be analyzed to a contractionary monetary shock, against a no-shock baseline. In addition, the results from both shock variables are compared

and tested for robustness.

IV.1 Local Jordá projection approach

To determine the impact of monetary shocks on the real economy, we compute the impulse-response function for various real-economic indicators. We follow Romer and Romer (2023) and use Jordá local projections (see, Jordá 2005). This method applies a sequence of linear regressions on the selected dependent variable Y , described in Equation (3). The equation is solved for each value of h representing different time horizons of the projection.

$$Y_{t+h} = \beta_0^h + \beta_1^h S_t + \sum_{k=1}^K \phi_k^h S_{t-k} + \sum_{k=1}^K \theta_k^h Y_{t-k} + e_t^h. \quad (3)$$

The variable Y represents the selected macroeconomic or financial variable, S represents the identified shock variable, and K is the number of lags utilized for both the result and shock variables. The lag and horizon variables are determined based on the methodology proposed by Romer and Romer (2023). For monthly data, a total of twelve delays were used, while for quarterly data, four lags were utilized for the variables. The projected time frame was either sixty months or twenty quarters resulting in a year horizon for both cases. The latencies of the outcome variables capture the typical patterns of the series, while the coefficients of the shock variables may detect the impacts of the monetary shocks, thus representing the response of the outcome variable to the monetary shocks Romer and Romer (2023).

IV.1.1 Defining the shock variable

In the local projection technique, the shock variable S may be represented as a binary variable. A value of 0 indicates the absence of a shock, while a value of 1 shows the presence of a shock at time t . Moreover, a more intricate approach involves utilizing the value of 1 to denote a contractionary shock, while -1 is used to signify an expansionary shock. This methodology is presented in section 2.1.1 and adopted by Romer and Romer (2023) as well. In addition, the paper suggests that monetary shocks are consistently present, varying in severity, and may be considered as a continuous variable. To measure such continuous shocks, the stance of monetary policy is used as the variable S . As inferred from Equation 1, the sign convention of the monetary stance score will

adhere to the traditional shock variables. Consequently, the local projection will assess the impact of a contractionary monetary shock on the economy.

IV.2 Response to monetary shocks

To evaluate the impact of a monetary shock on real economic and financial variables, we use the local Jordá projection method, which was explained in a previous section. The response functions are computed for a total of four variables: unemployment, real GDP, and two inflation measures, namely CPI and core inflation. Initially, the conventional shock measure is used as the independent variable and then contrasted with the measure generated by the RoBERTa fine-tuned and Chat GPT models. The response function is graphed and analyzed, and then the shock measurements are compared.

IV.2.1 Unemployment

In order to analyze the response function of monthly unemployment, the percentage of unemployed people compared to the total labor force was used. Figure 3 displays the outcomes, depicting the response of the unemployment rate in relation to a baseline without any monetary shocks. The Figure also shows the 95% confidence interval for the estimates. The projections are made 60 months or 5 years in advance. A contractionary monetary shock leads to an increase in unemployment in subsequent periods. This increase is substantial, lasting around one year and then diminishing within two years. While the same general conclusion can be said about the response to the classic shock variable, the effect is never significant, except after three and a half years. Also, the shock doesn't seem to disappear, and unemployment doesn't drop back to zero after sixty periods. Furthermore, the effects measured are much smaller, only amounting to about a 0.6 percentage point increase at the peaks, while both text-analysis-based models suggest an increase of around 1.5 to 1.8 percentage points.

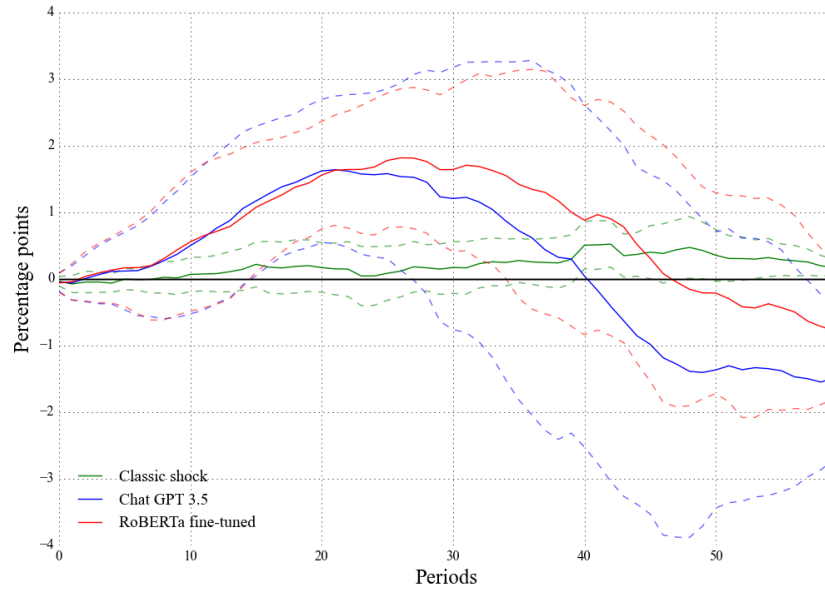


Figure 4: Response of unemployment to a contractionary monetary shock

Notes: The figure shows the response of the monthly unemployment rate to a contractionary monetary shock, projected out sixty months or five years. The results are plotted against a no-shock baseline, in percentage points. The projections were calculated using local Jordá projections, using twelve lags for both unemployment and shock variables. In the plot, the projections are based on two text-analysis-based shocks and shock measured in a classical way. The continuous lines show the mean projections, with two 95% confidence interval bands around them.

This phenomenon might be attributed to the increase in borrowing costs, which hampers investments and thus decelerates economic development. This would restrict the expenditure options of enterprises, leading to increased hesitancy in hiring new personnel or possibly initiating layoffs, both contributing to an increase in unemployment.

IV.2.2 Real GDP

In order to assess the impact of a monetary shock on the real GDP, we computed the logarithms of the quarterly series and regressed on our shock measures. Figure 4 displays the reaction function to contractionary monetary shocks as assessed using both our text analysis method and the conventional methodology. In addition to the forecasts, the 95% confidence interval is shown. The outcomes are graphed in comparison to a baseline scenario without any shocks and are predicted over a period of twenty quarters or five years.

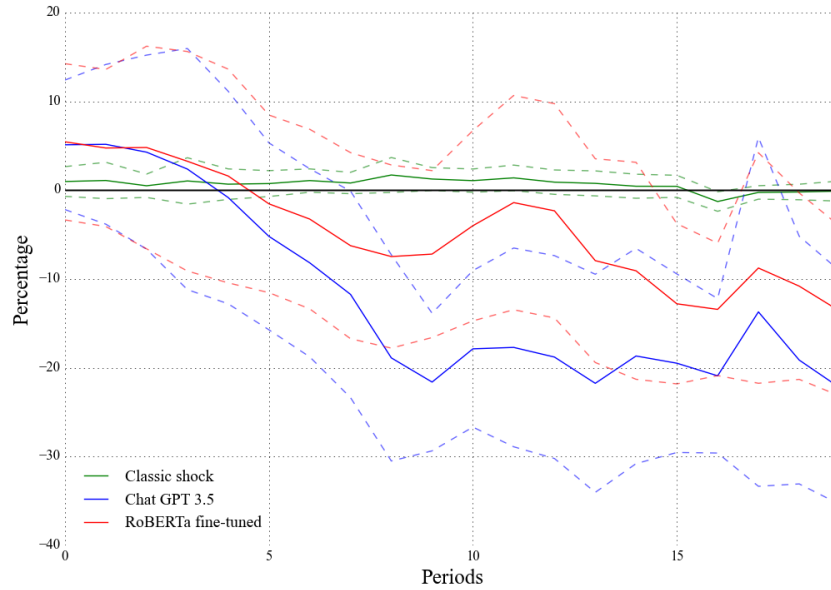


Figure 5: Response of real GDP to a contractionary monetary shock

Notes: The figure shows the response of the quarterly real GDP to a contractionary monetary shock, projected out twenty quarters or five years. The results are plotted against a no-shock baseline, in percentages. The projections were calculated using local Jordá projections, using four lags for both real GDP and shock variables. In the plot, the projections are based on two text-analysis-based shocks and shock measured in a classical way. The continuous lines show the mean projections, with two 95% confidence interval bands around them.

The text-analysis-based shock measurements show that the real GDP decreases in reaction to a contractionary shock, albeit the magnitudes of the decrease vary significantly. While most measures are not significant for up to four years, the models suggest that GDP declines heavily, up to -13% (RoBERTa) or -22% (Chat GPT). After a duration of four years, this reduction reaches a level of significance at a rate of 5%. In response to the classic shock measure the real GDP experiences a period of stagnation followed by a decline of 1% after four years. Overall, it can be inferred that the Hungarian economy's production would decrease as a result of a contractionary monetary shock, regardless of the specific shock variable utilized in the projection calculations. This conclusion aligns with conventional economic theory since a contractionary shock leads to higher borrowing costs and increased caution among both private and public sector institutions. Consequently, investments by both sectors are delayed, resulting in a negative impact on economic production, as measured by real GDP. Furthermore, the findings may indicate the presence of

omitted variable bias in the estimate, which distorts the results away from zero. An example would be the presence of high-pressure economy maintained by policymakers during the investigated time period. Additionally, the choice of a short time frame, during which many recessionary episodes occurred in Hungary, may also contribute to the skewed results.⁴

IV.2.3 Inflation

Finally, we measure the impact of a contractionary monetary shock on inflation levels. For the analysis, both CPI and core inflation are used, at quarterly intervals. The differences in the logarithm of both variables were calculated and regressed upon the chosen shock variables. Figure 6 shows the results for CPI and Figure A.4 shows displays the reaction of core inflation, against a no-shock baseline. All projections are calculated twenty quarters in advance and are plotted with their 95% confidence interval.

The results obtained through the text-analysis-based shock are not reflecting economic theory, nor empirical evidence found before. Although a rise in inflation is not unusual immediately following a monetary shock, after a time period, the shock should cause inflation to decline (Romer and Romer, 2023). While the response functions declined after 4 quarters and became negative for two quarters, in the case of the RoBERTa, it rises rapidly again. However, it is important to mention that there are very few significant periods, other than the initial rise following the shock. Contrarily, the classical shock measure prints a much clearer picture, where initial higher inflation will dissipate and turn negative after six quarters in the case of CPI and eight quarters in the case of core inflation, driving inflation down by 0.5%. These results suggest that the monetary transmission time in Hungary is around one and a half to two years, and affects the most volatile assets first, found in CPI.

⁴A possible future research may want to examine this issue using more robust statistical frameworks, and a longer time period.

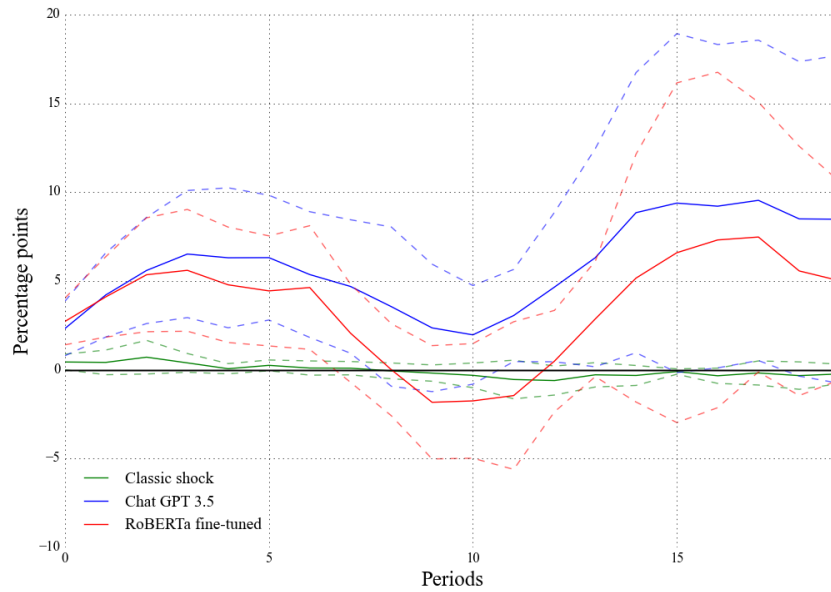


Figure 6: Response of CPI to a contractionary monetary shock

Notes: The figure shows the response of the quarterly CPI to a contractionary monetary shock, projected out twenty quarters or five years. The results are plotted against a no-shock baseline, in percentage points. The projections were calculated using local Jordá projections, using four lags for both CPI and shock variables. In the plot, the projections are based on two text-analysis-based shocks and shock measured in a classical way. The continuous lines show the mean projections, with two 95% confidence interval bands around them.

A possible explanation for the unusual findings is that the last data point of inflation is from 2023Q2, where inflation was still significantly high in Hungary, and the first-rate hikes by the MNB started in mid-2021 and the first increase of 100 basis points happened in the beginning of 2022. Thus, our measured monetary transmission times have not been sufficient to have a decreasing impact on inflation. This may introduce bias into our data, and in order to eliminate it, the same functions were calculated on a sample ranging from 2008 to 2020. Figures 7 and A.6 show the response functions of CPI and core inflation.

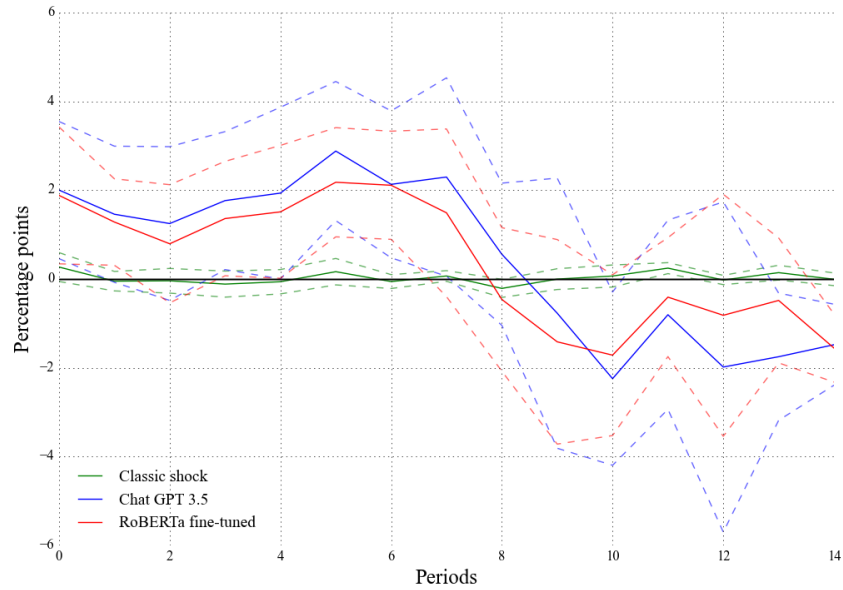


Figure 7: Response of CPI to a contractionary monetary shock (reduced sample, ranging 2008-2020)

Notes: The figure shows the response of the quarterly CPI to a contractionary monetary shock, on a reduced sample ranging from 2008 to 2020, projected out twenty quarters or five years. The results are plotted against a no-shock baseline, in percentage points. The projections were calculated using local Jordá projections, using four lags for both CPI and shock variables. In the plot, the projections are based on two text-analysis-based shocks and shock measured in a classical way. The continuous lines show the mean projections, with two 95% confidence interval bands around them.

When using a smaller sample size, the impacts become more evident, and after an initial period of elevated inflation, it gradually decreases after eight and nine quarters for the Consumer Price Index (CPI) and core inflation, respectively. While reducing the sample size created unbiased inflation response measurements, it also led to the exclusion of some classical shocks. As a result, our estimate, based on classical shocks, has become volatile and lacks significance for the majority of data points. However, after eight quarters, the response to the classic shock projection shows significant negative results. To summarize, a contractionary monetary shock leads to an immediate increase in inflation that starts to decrease after four to six quarters. Additionally, following a period of monetary transmission, typically between six and nine quarters according to our calculations, both measures of inflation begin to diminish and turn negative. The reason for this phenomenon is the deliberate deceleration of the economy by the MNB, resulting in stagnant growth, which in turn leads to simultaneous stagnation in both prices and wages.

IV.2.4 Robustness

To assess the robustness of our findings, we computed the results using all four of our sentiment metrics. Similar findings were obtained for unemployment, real GDP, and inflation. However, the SetFit model exhibited instability in all calculations, displaying significant volatility. However, the findings from all metrics consistently reached the same conclusion, so demonstrating the strength and reliability of our study. The Appendix has comprehensive charts displaying response functions for all measurements.

V Conclusions

This research used language modeling to identify monetary shocks by analyzing the Hungarian Central Bank's press releases. Using the acquired metrics, a shock analysis was undertaken, for effects on unemployment, real GDP, and inflation. Ultimately, our sentiment measure demonstrated significance as shown by the correlation analysis and the Granger causality tests. The monetary stance serves as a predictor for several macroeconomic and financial factors, such as the nominal interest rate determined by the MNB, the Euro-Forint exchange rate, unemployment, real GDP, and inflation. Moreover, the shock analysis demonstrated the efficacy of our sentiment measures compared to the traditional shock measures, as it yielded more robust and statistically significant outcomes overall. Based on the results of this study, monetary policy indeed has an impact on the real economy. A unit-size contractionary monetary shock has a negative impact on both unemployment and real GDP, while it reduces inflation, but it takes between 6 to 9 quarters for the effects of such shocks to be transmitted. Furthermore, it can be confidently said that our findings demonstrated robustness since consistent outcomes were obtained utilizing any of our sentiment metrics.

References

- P. Adämmmer. lpirfs: An R package to estimate impulse response functions by local projections. *The R Journal*, 11(2):421, 2019. URL <https://doi.org/10.32614/rj-2019-052>.
- D. Araci. FinBERT: Financial sentiment analysis with pre-trained language models. *arXiv*, 2019. URL <https://doi.org/10.48550/arxiv.1908.10063>.
- H. C. Bjørnland. Monetary policy and exchange rate interactions in a small open economy. *Scandinavian Journal of Economics*, 110(1):197–221, 2008. doi: 10.1111/j.1467-9442.2008.00532.x. URL <https://doi.org/10.1111/j.1467-9442.2008.00532.x>.
- A. S. Blinder, M. Ehrmann, M. Fratzscher, J. De Haan, and D.-J. Jansen. Central bank communication and monetary policy: A survey of theory and evidence. *Journal of Economic Literature*, 46(4):910–945, 2008. URL <https://doi.org/10.1257/jel.46.4.910>.
- A. S. Blinder, M. Ehrmann, J. De Haan, and D.-J. Jansen. Central bank communication with the general public: Promise or false hope? 2022. URL <https://doi.org/10.3386/w30277>.
- T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, J. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, J. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei. Language models are few-shot learners. *arXiv*, 2020. URL <https://arxiv.org/pdf/2005.14165.pdf>.
- J. De Haan and J.-E. Sturm. Central bank communication. pages 230–262, 2019. URL <https://doi.org/10.1093/oxfordhb/9780190626198.013.8>.
- J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv*, 2018. URL <https://doi.org/10.48550/arxiv.1810.04805>.
- Eurostat. Unemployment by sex and age – monthly data. 2024. https://ec.europa.eu/eurostat/databrowser/view/UNE_RT_M__custom_10742670/default/\table.

- T. Evdokimova, P. N. Mohacsi, O. Ponomarenko, and E. Ribakova. Overtaking the masters? comparison of policy communication of emerging market central banks with the fed and ecb. *PIIE Working Paper*, 2023.
- J. Galí and T. Monacelli. Monetary policy and exchange rate volatility in a small open economy. *Review of Economic Studies*, 72(3):707–734, 2005. doi: 10.1111/j.1467-937x.2005.00349.x. URL <https://doi.org/10.1111/j.1467-937x.2005.00349.x>.
- Y. Gorodnichenko, T. Pham, and O. Talavera. The voice of monetary policy. 2021. URL <https://doi.org/10.3386/w28592>.
- C. W. J. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424, 1969. URL <https://doi.org/10.2307/1912791>.
- P. Gábel and K. Pintér. The effect of the mnb’s communication on financial markets. Working Paper 2006/9, Magyar Nemzeti Bank, Budapest, 2006.
- Ó. Jordá. Estimation and inference of impulse responses by local projections. *The American Economic Review*, 95(1):161–182, 2005. URL <https://doi.org/10.1257/0002828053828518>.
- A. Koubaa. GPT-4 vs. GPT-3.5: A concise showdown. 2023. doi: 10.36227/techrxiv.22312330.v2.
- KSH. *A Fogyasztóiár-statisztika módszere*. 2000.
- Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov. ROBERTA: A robustly optimized BERT pretraining approach. *arXiv*, 2019. URL <https://doi.org/10.48550/arxiv.1907.11692>.
- L. Loukas, I. Stogiannidis, P. Malakasiotis, and S. Vassos. Breaking the bank with chatGPT: Few-shot text classification for finance. *arXiv*, 2023. URL <https://doi.org/10.48550/arxiv.2308.14634>.
- MNB. Inflációs jelentés előrejelzési adatbázis. Dataset, 2024.

- OpenAI. GPT-4 technical report. *arXiv*, 2023. URL <https://doi.org/10.48550/arxiv.2303.08774>.
- C. D. Romer and D. Romer. Presidential address: Does monetary policy matter? the narrative approach after 35 years. *The American Economic Review*, 113(6):1395–1423, 2023. URL <https://doi.org/10.1257/aer.113.6.1395>.
- M. Rozkrut, K. Rybiński, L. Sztaba, and R. Szwaja. Quest for central bank communication: Does it pay to be “talkative”? *European Journal of Political Economy/Europäische Zeitschrift Für Politische Ökonomie*, 23(1):176–206, 2007. doi: 10.1016/j.ejpoleco.2006.09.011. URL <https://doi.org/10.1016/j.ejpoleco.2006.09.011>.
- A. Seth. Granger causality. *Scholarpedia*, 2(7):1667, 2007.
- A. Shah and S. Chava. Zero is not hero yet: Benchmarking zero-shot performance of LLMs for financial tasks. *Social Science Research Network*, 2023. URL <https://doi.org/10.2139/ssrn.4458613>.
- A. Shah, S. Paturi, and S. Chava. Trillion dollar words: a new financial dataset, task & market analysis. *arXiv*, 2023. URL <https://doi.org/10.48550/arxiv.2305.07972>.
- A. Shojaie and E. B. Fox. Granger causality: a review and recent advances. *Annual Review of Statistics and Its Application*, 9(1):289–319, 2022. doi: 10.1146/annurev-statistics-040120-010930. URL <https://doi.org/10.1146/annurev-statistics-040120-010930>.
- C. A. Sims. Macroeconomics and reality. *Econometrica*, 48(1):1, 1980. URL <https://doi.org/10.2307/1912017>.
- D. I. Stern. From correlation to granger causality. *Social Science Research Network*, 2011. doi: 10.2139/ssrn.1959624. URL <https://doi.org/10.2139/ssrn.1959624>.
- L. C. Tunstall, N. Reimers, U. E. S. Jo, L. Bates, D. Korat, M. Wasserblat, and O. Pereg. Efficient few-shot learning without prompts. *arXiv*, 2022. URL <https://doi.org/10.48550/arxiv.2209.11055>.

J. Ye, X. Chen, N. Xu, C. Zu, Z. Shao, S. Liu, Y. Cui, Z. Zhou, C. Gong, Y. Shen, J. Zhou, S. Chen, T. Gui, Q. Zhang, and X. Huang. A comprehensive capability analysis of GPT-3 and GPT-3.5 series models. *arXiv*, 2023. URL <https://doi.org/10.48550/arxiv.2303.10420>.

A Appendix

A.1 Figures

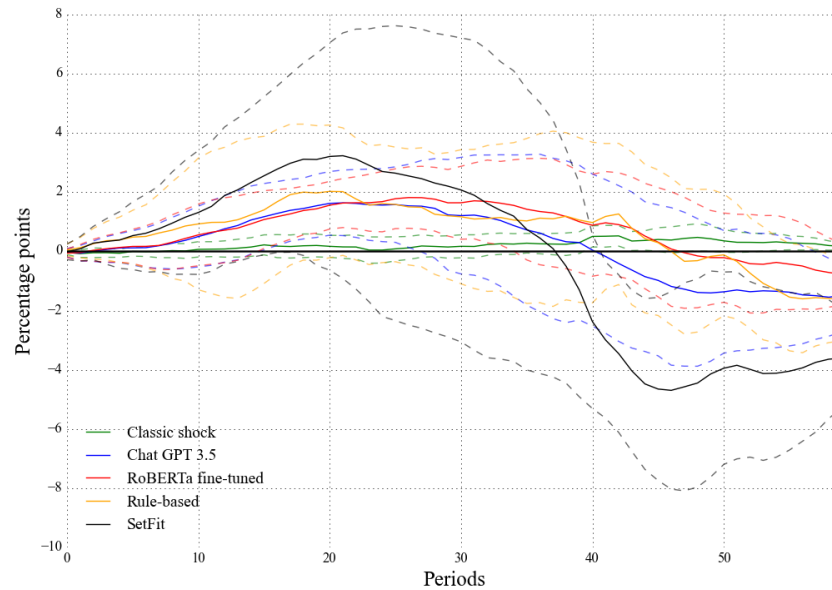


Figure A.1: Response of unemployment to a contractionary monetary shock, all measures

Notes: The figure shows the response of the monthly unemployment to a contractionary monetary shock, projected out sixty months or five years using all monetary shock measures. The results are plotted against a no-shock baseline, in percentage points. The projections were calculated using local Jordá projections, using twelve lags for both unemployment and shock variables. In the plot, the projections are based on the four text-analysis-based shocks and shock measured in a classical way. The continuous lines show the mean projections, with two 95% confidence interval bands around them.

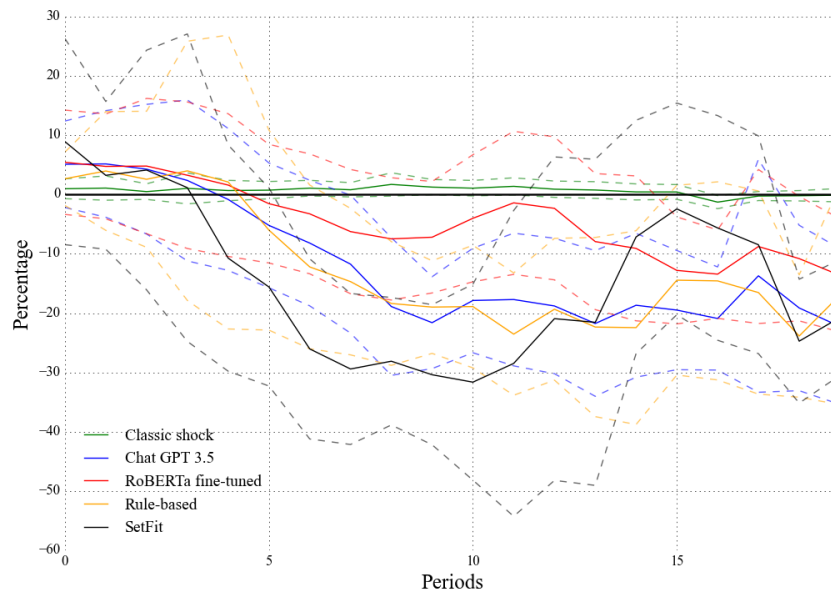


Figure A.2: Response of real GDP to a contractionary monetary shock, all measures

Notes: The figure shows the response of the quarterly real GDP to a contractionary monetary shock, projected out twenty quarters or five years, using all monetary shock measures. The results are plotted against a no-shock baseline, in percentages. The projections were calculated using local Jordá projections, using four lags for both real GDP and shock variables. In the plot, the projections are based on the four text-analysis-based shocks and shock measured in a classical way. The continuous lines show the mean projections, with two 95% confidence interval bands around them.

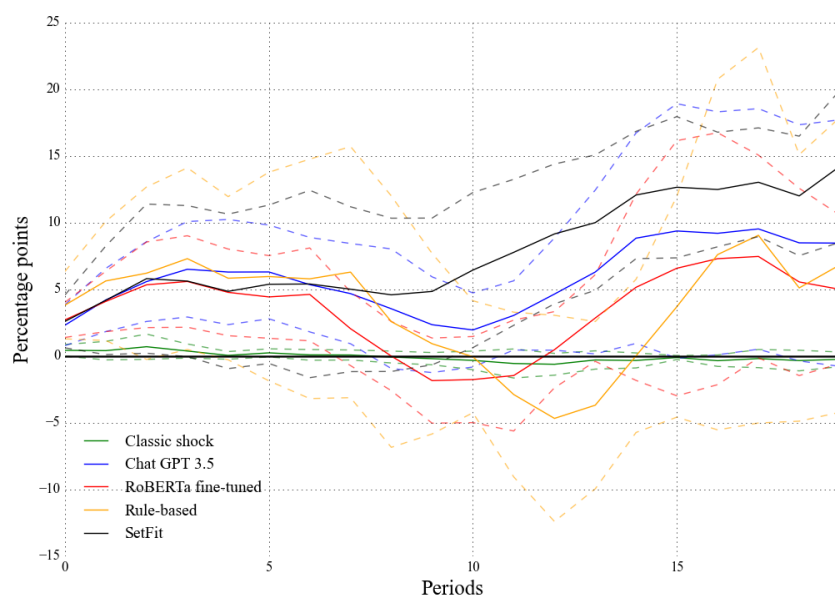


Figure A.3: Response of CPI to a contractionary monetary shock, all measures

Notes: The figure shows the response of the quarterly CPI to a contractionary monetary shock, projected out twenty quarters or five years using all monetary shock measures. The results are plotted against a no-shock baseline, in percentage points. The projections were calculated using local Jordá projections, using four lags for both CPI and shock variables. In the plot, the projections are based on the four text-analysis-based shocks and shock measured in a classical way. The continuous lines show the mean projections, with two 95% confidence interval bands around them.

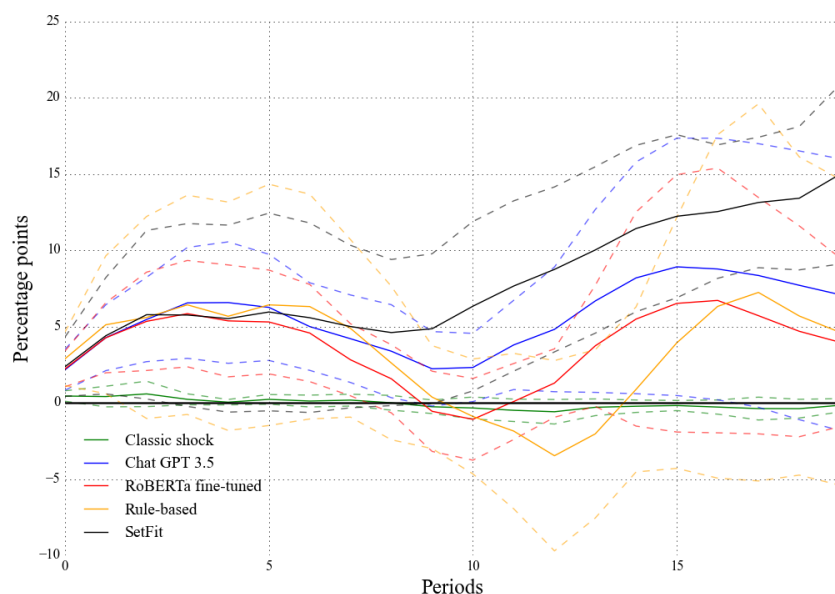


Figure A.4: Response of core inflation to a contractionary monetary shock, all measures

Notes: The figure shows the response of the quarterly core inflation to a contractionary monetary shock, projected out twenty quarters or five years using all monetary shock measures. The results are plotted against a no-shock baseline, in percentage points. The projections were calculated using local Jordá projections, using four lags for both core inflation and shock variables. In the plot, the projections are based on the four text-analysis-based shocks and shock measured in a classical way. The continuous lines show the mean projections, with two 95% confidence interval bands around them.

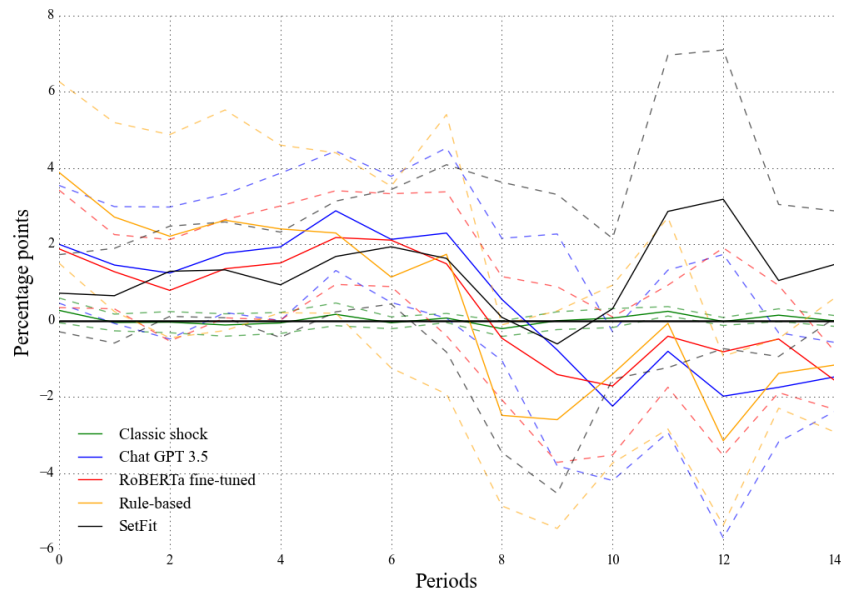


Figure A.5: Response of CPI to a contractionary monetary shock (reduced sample, ranging 2008-2020), all measures

Notes: The figure shows the response of the quarterly CPI to a contractionary monetary shock, on a reduced sample ranging from 2008 to 2020, projected out twenty quarters or five years using all monetary shock measures. The results are plotted against a no-shock baseline, in percentage points. The projections were calculated using local Jordá projections, using four lags for both CPI and shock variables. In the plot, the projections are based on the four text-analysis-based shocks and shock measured in a classical way. The continuous lines show the mean projections, with two 95% confidence interval bands around them.

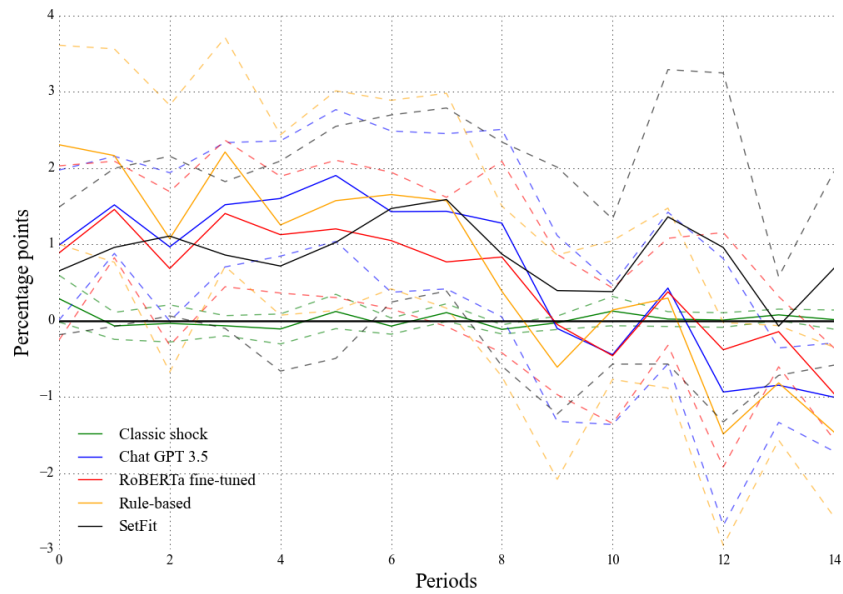


Figure A.6: Response of core inflation to a contractionary monetary shock (reduced sample, ranging 2008-2020), all measures

Notes: The figure shows the response of the quarterly core inflation to a contractionary monetary shock, on a reduced sample ranging from 2008 to 2020, projected out twenty quarters or five years using all monetary shock measures. The results are plotted against a no-shock baseline, in percentage points. The projections were calculated using local Jordá projections, using four lags for both core inflation and shock variables. In the plot, the projections are based on the four text-analysis-based shocks and shock measured in a classical way. The continuous lines show the mean projections, with two 95% confidence interval bands around them.

A.2 Tables

Panel A1	Panel A2
inflation expectation, interest rate, bank rate, fund rate, price, economic activity, inflation, employment	anchor, cut, subdue, decline, decrease, reduce, low, drop, fall, fell, decelerate, slow, pause, pausing, stable, non-accelerating, downward, tighten
Panel B1	Panel B2
unemployment, growth, exchange rate, productivity, deficit, demand, job market, monetary policy	ease, easing, rise, rising, increase, expand, improve, strong, upward, raise, high, rapid
Panel C	
weren't, were not, wasn't, was not, did not, didn't, do not, don't, will not, won't	

Table A.1: Words/phrases to classify monetary text

Notes: This table shows the words and phrases needed to classify monetary texts into tightening or easing sentiment (Gorodnichenko et al., 2021). A sentence is considered to be related to monetary easing if it contains a word from A1 and A2 or B1 and B2. If it contains words from A1 and B2 or A2 and B1, it is considered a tightening sentiment. In the case a sentence contains any word from C, the classification is reversed (Shah et al., 2023).

Variables	Chat GPT 3.5	RoBERTa fine-tuned	SetFit	Rule-based	Classic shock	RGDP	CPI	CCP
Chat GPT 3.5	1.000	.	.	.	.	.	.	.
RoBERTa	0.909	1.000	.	.	.	.	.	.
SetFit	0.783	0.698	1.000	.	.	.	.	.
Rule-based	0.621	0.577	0.450	1.000	.	.	.	.
Classic shock	0.211	0.253	0.091	0.200	1.000	.	.	.
RGDP	0.711	0.612	0.841	0.195	0.101	1.000	.	.
CPI	0.528	0.467	0.576	0.060	0.129	0.835	1.000	.
CCP	0.530	0.473	0.591	0.053	0.133	0.847	0.996	1.000

Table A.2: Correlation matrix for variables with quarterly frequencies

Notes: Shock variables are averaged across the three-month periods.

Variables	Chat GPT 3.5	RoBERTa fine-tuned	SetFit	Rule-based	Classic shock	EURHUF	Unemployment	Base rate change
Chat GPT 3.5	1.000	.	.	.	.	.	.	.
RoBERTa	0.829	1.000	.	.	.	.	.	.
SetFit	0.726	0.641	1.000	.	.	.	.	.
Rule-based	0.477	0.462	0.313	1.000	.	.	.	.
Classic shock	0.131	0.124	0.047	0.074	1.000	.	.	.
EURHUF	0.456	0.357	0.547	-0.047	0.039	1.000	.	.
Unemployment	-0.464	-0.385	-0.716	0.024	0.007	-0.712	1.000	.
Base rate difference	0.513	0.491	0.368	0.278	0.012	0.203	-0.229	1.000

Table A.3: Correlation matrix for variables with monthly frequency