Big Data and Inflation Forecasting in Nigeria: A Text Mining Application[1]

Adebiyi, M. A., Adenuga, A. O., Olusegun, T. S., and Mbutor, O. O.*

Abstract

The success of monetary policy is substantially predicated on the availability of reliable forecast of inflation. However, the shocks arising from COVID-19 and the Russia-Ukraine war have brought about significant economic uncertainties; thus, necessitating the fine-tuning of existing forecasting models of the Central Bank of Nigeria. This study explores the usefulness of public sentiments obtained using machine learning methods to improve the predictive power of the existing short-term inflation forecasting model (STIF) in Nigeria. Findings indicate that, for all components of inflation, models that include the computed sentiment index perform better in both in-sample and out-sample forecasts than those excluding the index. Thus, we conclude that sentiment-based inflation forecasting models are useful for improving the headline inflation forecast and suggest the use of forward guidance monetary instruments in the form of "open mouth operations", to ensure economic agents' sentiments are well anchored.

Keywords: Big data, Forecast, Inflation
JEL Classification : C53, C55, E31

I. Introduction

Economic forecasts are fundamental to the design and implementation of economic policies. Policy institutions, especially central banks, rely on inflation forecasts to inform monetary policy decisions and anchor inflation expectations. Thus, improving the accuracy of economic forecasts remains an integral part of effective policymaking. One factor which affects the precision of economic forecasts is the choice of variables included in economic models. It is therefore important that model variables are well curated to optimise forecast outcomes. Over the years, several efforts have been made to develop economic indices that help to better predict the behaviour of economic indicators, including indices of economic policy uncertainty and text-based sentiments indicators.

---

* *Adebiyi, M. A. is the Director, Research Department, while Adenuga, A. O., Olusegun, T. S., and Mbutor, O. O. are staff of the Department, Central Bank of Nigeria. The views expressed in this paper are those of the authors and do not necessarily reflect the opinions of the Central Bank of Nigeria.*

[1]*We are sincerely grateful to the contributions and technical assistance provided by Osaretin Evbuomwan, Idoko A. Itodo, Obioma C. Asuzu, Nwamaka G. Nwafor, Tony A. Odu, Nuhu Ado and Umar Yusuf, respectively. They are staff of the Macroeconomic Modeling Division, Research Department, Central Bank of Nigeria, Abuja. The comments of anonymous reviewers are gratefully acknowledged.*

The CBN houses a suite of macro-econometric models that have been used to forecast inflation over the years. These include the Inflation Forecasting Model for Nigeria (Adebiyi et al., 2010), the Dynamics of Inflation in Nigeria (Mordi et al., 2012) and Short-Term Inflation Forecasting (STIF) model (Mordi et al., 2013); among others; with the STIF model being the benchmark model for forecasting inflation in Nigeria currently. Although these models have proved reliable in predicting inflation over time, the changing economic landscape, arising from COVID-19 shocks and the economic uncertainties brought about by the Russia-Ukraine war, have further complicated the economic and policy environments. It is in the light of these developments that a review and possible fine tuning of existing inflation forecasting models is of significant policy imperatives.

This paper explores the usefulness of public sentiments obtained using machine learning techniques in improving the extant inflation forecasting models of the Bank. We build on the STIF of the CBN, incorporating a sentiment index constructed with data on public perception about inflation, obtained from social media posts, particularly Twitter. The sentiments are generated from the 'opinions' expressed about price developments and inflation expectations in the Nigerian economy – an inflation perceptions index.

Studies have provided evidence of forecast improvements following the incorporation of sentiment-based indices obtained from media data (Tetlock, 2007) while analysing trading volume (Jegadeesh & Wu, 2013) and stock returns (Kearney & Liu, 2014). Also, Ardia et al. (2019) improved the economic growth forecasts for the US by including computed sentiment indices in their econometric models.

Our findings show that inflation forecasts improved by including an inflation perception index generated from sentiments expressed by the public on social media platforms. From an ex-post sample forecast analysis, we found that inflation predictions from the short-term inflation forecast STIF-SA model outperformed the benchmark STIF model. The forecasts of inflation and its components using the sentiment index showed that core and food inflation are expected to rise slightly in August 2022 and September 2022. However, food inflation is expected to moderate slightly in October 2022, while the core is expected to rise marginally in the same month. The net impact on the headline is a marginal increase in August, September, and October 2022.

The remaining sections of the study are organised as follows: Section 2 presents the related theory and relevant literature. Section 3 focuses on stylised facts and inflation dynamics in Nigeria over the years; while Section 4, outlines the methodology, including details relating to the data used. Section 5 discusses and evaluates the results of the various models while Section 6 concludes the paper with a summary and some policy implications.

II. Literature Review

II.1 Theoretical Framework

Theories of inflation sentiments or expectations encompass the behaviour and perception of economic agents towards changes in the price level. Though the relationship between the effects of inflation and the economy as well as inflation's impacts on other macroeconomic variables are explained by other theories such as the quantity theory of money, real theory, structural theory, and fiscal theory of price level, this paper is hinged primarily on the cognitive narrative theory, which has been found to explain the role of big data in inflation forecasting (Ramey & Shapiro, 1999; Choi & Varian, 2012; Dominguez & Shapiro, 2013; Tuckett et al., 2015).

The cognitive narrative theory explains how social-psychological sentiments of economic agents influence macroeconomic outcomes (Nyman et al., 2014). It hinges on two emotions – "anxiety" about possible future loss and "excitement" about possible future profit. The theory explains that sentiments such as "positive or negative", "optimistic or pessimistic", are found in narrative texts written by rational economic agents (Chong et al., 2014). Although sentiments about economic variables such as inflation, GDP, and interest rate cannot be directly observed, they are estimated using sample survey data and, more recently, big data approaches. Recently, several important studies have adopted these techniques. Using machine learning algorithms, earlier researchers have presented studies that capture the relationship between inflation expectations expressed in public sentiments and macroeconomic variables. The underlying idea in most of these studies is that changes in information about the economy may alter expectations and, thus, economic behaviour (Tuckett et al., 2015).

II.2 Empirical Literature Review

Inflation is driven by internal and external factors to an economy (Adeleye et al., 2019; Forbes et al., 2018; Nagy & Tengely, 2018; Pop & Murărașu, 2018). The internal drivers of Inflation include the money supply, government expenditures, net food exports, domestic demand and supply, political instability or insecurity, and lending interest rates.

The external drivers of inflation are magnified by the increasing level of integration among countries into a global production process. This is inherent in the impact of international business cycles on inflation. As a result, exchange rate, imported inflation, and trade openness are the key external triggers of inflation (Rodriguez & Yoldas, 2016). Nagy and Tengely (2018) reveal that domestic inflation is more sensitive to global factors than domestic factors in Hungary. Similarly, Adeleye et al. (2019) disclose that inflation in Nigeria is driven majorly by external factors. In Kenya and Ethiopia, world food prices and exchange rates have a long-run impact, while money growth and agricultural supply shocks have short-to-medium-run effects on inflation (Durevall & Sjö, 2012).

Since the work of Keynes (1936), economists have acknowledged the role of expectations and sentiments in influencing economic decision-making (Algaba et al., 2020). However, literature on sentiment analysis and its role in determining macroeconomic outcomes is nascent in economics. Notwithstanding, scholars have recently given increasing attention to the role of public sentiment in predicting macroeconomic variables such as inflationary trends.

In addition, scholars employ sentiment analysis and other modern approaches based on machine learning techniques to improve the central banks' inflation forecast. Kohlscheen (2021) utilise machine learning techniques and finds that consumer expectations significantly influence inflation in advanced economies. Simionescu (2022) argues that sentiment analysis is a valuable technique to correctly discern public opinion on some topics of interest. It has also improved the forecast accuracy of existing inflation forecasting models, as they now outperform all the other traditional forecasts. This assertion corroborates the findings of Jones et al. (2020) and Xu et al. (2022). A possible reason for the success of the sentiment indices is due to its ability to forecast future uncertainties which arise due to structural changes, which point forecasts do not capture (Rambaccussing & Kwiatkowski, 2020). However, this differs remarkably from the finding of Clements and Reade (2020), which shows that whereas sentiment narratives could predict the errors in the numerical forecasts of output growth; there is no evidence that past changes in sentiment could predict subsequent changes in the point forecasts of Inflation.

Furthermore, Figueiredo (2010) finds that models that utilise big data have proven to outperform other traditional time series models in forecasting inflation. FAVAR, Bayesian regression, DFM with Cumulative GARCH, and other big data accommodating models perform better than other traditional time series models like structural VAR, GARCH, AR(p), and AR(p)-GARCH (Banbura et al., 2014; Bordoloi et al., 2010; Alessi et al., 2009; De Mol et al., 2008; Bernanke et al., 2005).

Empirical evidence suggests that forecasting with the use of Google search data, pioneered by the works of Baker et al. (2016) outperforms both AR (1) models and survey-based predictors (Niesert, 2020). It has improved the mean absolute prediction error in the US inflation and consumption, UK housing market, Swedish private consumption, and German and Israeli unemployment (Guzman, 2011; McLaren & Shanbhogue, 2011; Lindberg, 2011; Askitas & Zimmermann, 2009;

Suchoy, 2009; Vosen & Schmidt, 2011). Simionescu (2022) contends that sentiment forecasts based on narratives in the official publications of central banks outperformed the numerical projections and a variety of combined forecasts, including those of the Central Bank of Romania. Similarly, De Caigny et al. (2020) shows that incorporating textual data in a Consumer Churn Prediction (CCP) model improves predictive performance.

Sentiments from the Federal Reserve economic forecasts is strongly correlated with future economic performance - positively correlated with GDP, and negatively correlated with unemployment and inflation (Sharpe et al., 2022). It predicts positive monetary policy (fed funds rate) surprises and higher stock returns up to four quarters ahead. In a related study, Rambaccussing and Kwiatkowski (2020) affirm that newspaper sentiment does not improve nowcasts or short-term inflation forecasts. It, however, improves the forecast of unemployment and output growth. Clements and Reade (2020) assert that sentiment derived from the narrative of the Bank of England's quarterly inflation reports can predict errors in the numerical forecasts of output growth, but not inflation. Also, past changes in sentiment predict subsequent changes in the point forecasts of output growth or Inflation.

Kraaijeveld and Smedt (2020) realise that Twitter sentiment has the power to predict the price returns of cryptocurrencies. Also, the introduction of the sentiment index improves the forecast performance of prices in the stock exchange market (Wojarnik, 2021; Li et al., 2020). Exploring which sentiment index type has more predictive power in forecasting stock market volatility, Liang et al. (2020) show that the daily social media and internet media news sentiment indices significantly impact stock market volatility, while the sentiment index built by the traditional newspaper has no impact.

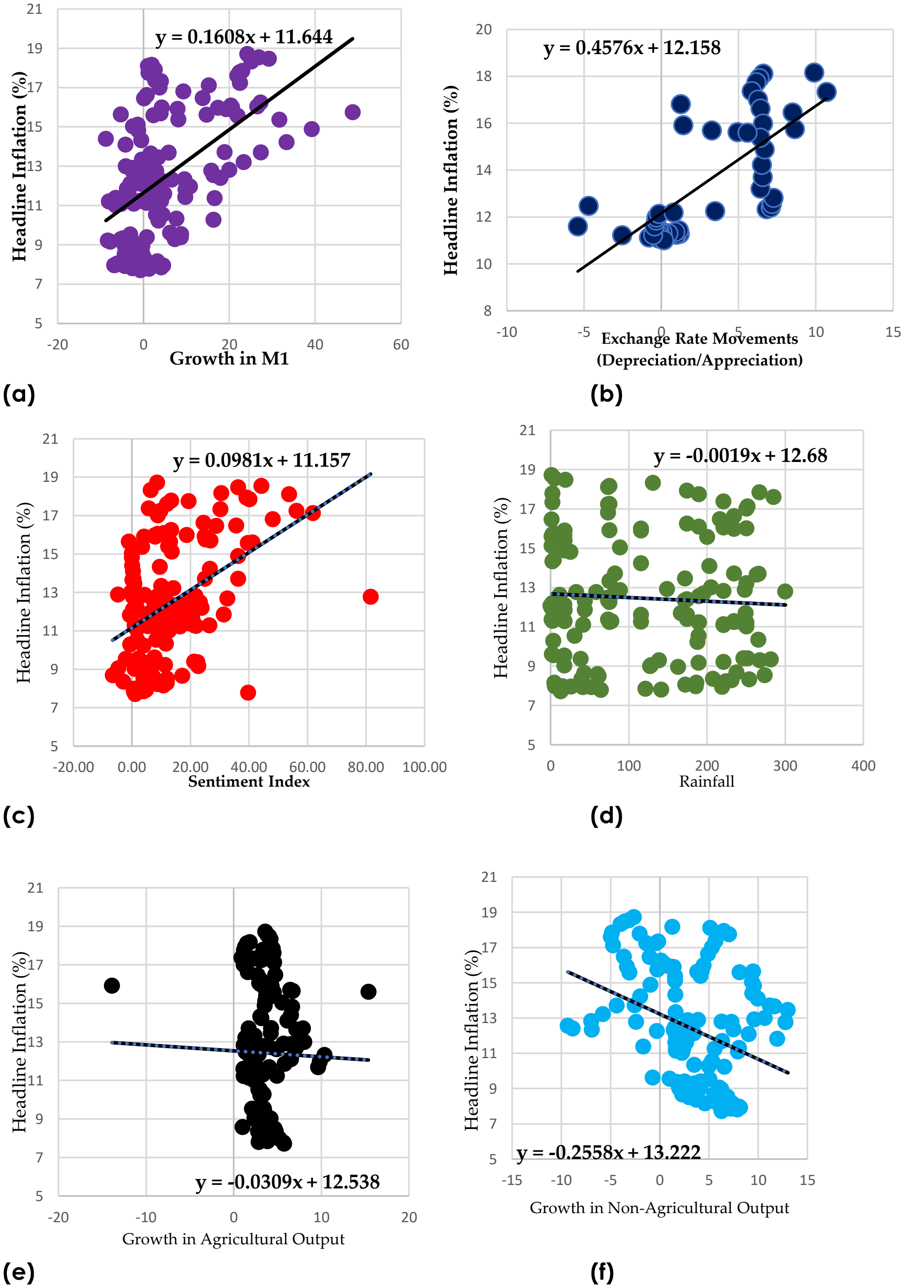
No previous attempt has been made to include sentiment indices in models used for forecasting inflation in Nigeria. Therefore, this study demonstrates the usefulness of social media sentiments in the operationalisation of the CBN's short term inflation forecasting model. The forecast evaluations are conducted for three components of inflation - food, core, and headline.

III. Stylised Facts

In Figures 1(a) – 1(f), we present scatter plots showing the association between headline inflation and selected variables, over the period 2010 – 2021. The selected variables are major drivers of inflation in Nigeria and are included in the short-term inflation forecasting model (STIF) adopted for this study. There appears to be a positive association between headline inflation and narrow money, sentiments, and exchange rate. Specifically, periods of rising inflation are associated with expansion in narrow money, exchange rate depreciation, and higher positive sentiments regarding inflation. In some economies, such as US and European countries, where public sentiments are high, central bankers influence the direction of key macroeconomic indicators, such as inflation and interest rates, without changing the policy rate through "open mouth operations[2]". Open-mouth operations or policy announcements can cause rates to deviate temporarily from their expected levels.

[2]Open mouth operations are announcements made by a central bank to achieve desired changes in its variables of interest, such as interest rate and prices. Such announcements cause economic agents to react in a manner that drives interest rates without the central bank having to set the policy rate directly.

Figure 1: Association between Headline Inflation and Selected Variables

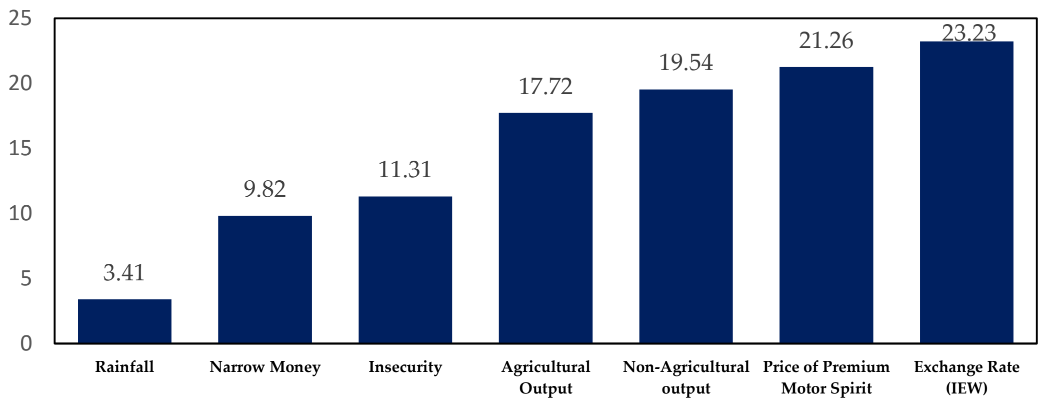


Source: Authors' computation.

III.1 Examining the Drivers of Inflation using Machine Learning

A stylised overview of drivers of inflation is presented in Figure 2. These include output in both the agricultural and non-agricultural sectors, import prices of commodities, such as energy and food and global economic downturn (ECB, 2017). Employing a random forest algorithm to identify inflation drivers in Nigeria over the period 2010:M01 – 2021:M12. The sample size was informed by data availability. Over the review period, the exchange rate was found to be the highest driver of inflation in Nigeria. This denotes that exchange rate pass-through is vital in influencing price variations. Other important determinants or drivers of inflation are PMI, agric and non-agric output, rainfall, and insecurity.

Figure 2: Drivers of Inflation and their Magnitudes (Full Sample 2010:01 – 2021:M12)



Source: Authors' computation.

IV. Methodology

IV.1 Data

In line with the variables included in the CBN's STIF model, the study utilises monthly data spanning 2010M1 - 2022M6 on headline, core, and food inflation; the annual growth rate of the narrow money (M1G); the manufacturing purchasers' managers index (PMI); non-manufacturing purchasers' managers index (NPMI); price of premium motor spirit (PMS); and exchange rate at the importers and exporters window (IEW). Other variables considered are real agricultural output (ARY); real non-Agricultural output; and real government expenditure, which includes both total recurrent and capital expenditure, as well as three independent indices. Data on the consumer price indices (CPI) for the headline, core, and food were sourced from the National Bureau of Statistics database while the sentiment index was computed using a Natural Language Processing (NLP). Data on the other variables were sourced from the statistical database of the CBN.

Using Twitter data, we construct the Sentiment index to capture news, individual sentiments, and market expectations. The collected unstructured tweets data are converted into structured data using a Natural Language Processing (NLP) system in Python software that classifies the polarity of each tweet into positive, negative, or neutral after assigning values to each tweet. It is a sentiment index that measures the public perception of the inflation dynamics and its drivers in Nigeria. The last two indices, referred to as the monetary index for agricultural sector growth, are labelled M_Index_ARY, and the monetary index for non-agricultural sector growth, M_Index_NARY. The monetary index for the agricultural output, M_Index_ARY, is the ratio of growth of real agric GDP (ARYG) to narrow money growth (M1G). This index indicates the level of agricultural output's monetary absorption. Similarly, the monetary index for the non-agric output.

The paper adopts an analytic approach similar to Clements and Reade (2020), Kraaijeveld and De Smedt (2020), and Xu et al. (2022). The extracted tweets are first subjected to a series of pre-processing steps before scores are generated using Python's Natural Language Toolkit (NLTK) library.

IV.2 Method of Data Analysis

Two main estimation techniques were employed in modelling the inflation dynamics in Nigeria. These techniques include the predictive regression (PR) model and Seasonal Autoregressive Integrated Moving Average Model with exogenous variables (SARIMA-X). The PR model was used in modelling and forecasting both food and core inflation, while the SARIMA-X was used for modelling and forecasting some of the drivers of food and core inflation, to enable the process of generating out-of-sample forecasts and simulations.

IV.2.1 The Predictive Regression Model

The framework of the STIF model is designed based on a predictive regression model, as outlined in Tule et al. (2019) and Westerlund and Narayan (2012). This model, which utilises the ordinary least squares (OLS) technique in its estimation, is of the following form:

$$food\_inf_t = \alpha + \beta Y_{t-1} + u_{1,t} \tag{1}$$

$$core\_inf_t = \varphi + pX_{t-1} + u_{2,t} \tag{2}$$

$$Headline\_inf_t = w_1 * food\_inf_t + w_2 * core\_inf_t \tag{3}$$

Where, $food\_inf_t$ is food inflation, while Y_{t-1} is a vector representing the one-period lag of the identified drivers of food inflation and $core\_inf_t$ is core inflation, while vector X_{t-1} is the one-period lag of the drivers of core inflation in Nigeria. Equation (3) is an identity, which expresses headline inflation as a weighted sum of the food and core inflation. The respective weights of food and core inflation, w_1 and w_2 are consistent with those provided by NBS in the headline CPI basket.

In implementing the model, equations (1) and (2) were estimated using a mix of drivers of the dependent variable. Specifically, equation (1) was estimated as follows:

$$food\_inf_t = \alpha + \beta Y_{t-1} + u_{1,t} \tag{4}$$

$$food\_inf_t = \alpha + \beta Y_{t-1} + \beta_f Sentiment + u_{1,t} \tag{5}$$

Equation (2) was also estimated, first, using the STIF model's drivers of core inflation in Nigeria, including the sentiment index. In the second estimation stage, the core inflation was estimated using all the drivers less sentiment index, as specified below:

$$core\_inf_t = \varphi + pX_{t-1} + p_c Sentiment + u_{2,t} \tag{6}$$

$$core\_inf_t = \varphi + pX_{t-1} + u_{2,t} \tag{7}$$

The in-sample forecasts of core and food inflation were subsequently generated from these models, and the performances of the models were evaluated using the mean square errors (MSEs). The idea is to determine whether including the sentiment index in both food and core inflation improves the respective forecasts of these two components of the headline inflation.

IV.2.2 The SARIMA Model

The seasonal autoregressive integrated moving average (SARIMA) model follows the methodology outlined by Box and Jenkins (1979). The first stage is to identify the orders of the $SARIMA(p, d, q) \times (P, D, Q)_s$, the second stage is to estimate the model, and the third stage is to perform diagnostic checks on the residuals to evaluate the fitted model. This process is repeated if the model selected fails diagnostic checks. The values of d and D (integration orders) are chosen so that the series become stationary, and after that, the SARIMA models are estimated using maximum likelihood approach.

IV.3 Implementation of the Model for Inflation (Model Evaluation Strategy)

The study performed both in-sample (2018M1 – 2022M6) and out-of-sample (2021M12 - 2022M6) forecasts of food and core variants of inflation. In each case, the MSE is evaluated to assess the contribution of the computed index of public sentiment on price developments in forecasting both food and core components of inflation. The out-of-sample forecast for headline inflation is obtained based on the forecasts of food and core components of inflation. The forecasts for food and core inflation included in the headline inflation forecasts are selected based on the MSE criterion in both in-sample and out-of-sample forecasts.

V. Results

V.1 Model Results

V.1.1 Food Inflation Models

Table 1 displays the parameter estimates of the food inflation model, with and without the public sentiment variable. Food inflation is expressed as a function of its past values (FOOD_INF), narrow money growth (M1G) lagged three months, non-manufacturing PMI (NPMI), rainfall (RAINFALL), and the lag of rainfall (RAINFALL (-1)) including some seasonal and trend variables.

Table 1: The Estimated Food Inflation Models

	Model A: Without Sentiment		**Model B: With Sentiment**	
Variable	**Coefficient**	**Prob.**	**Coefficient**	**Prob.**
FOOD_INF (-1)	0.92	0.00	0.89	0.00
M1G (-3)	0.03	0.00	0.03	0.00
LOG(NPMI)	0.61	0.08	0.61	0.08
LOG(RAINFALL)	-0.09	0.04	-0.08	0.10
LOG (RAINFALL (-1))	0.12	0.01	0.10	0.06
@SEAS (12)	0.34	0.11	0.35	0.10
@TREND	-0.61	0.00	-0.55	0.00
@TREND^2	0.00	0.00	0.00	0.00
@TREND^3	0.00	0.00	0.00	0.00
SENTIMENT			0.01	0.17

Source: Authors' estimate.

The coefficient estimates conform to a priori expectations and are correctly signed and significant at 10.0 per cent significance level for both models (with and without sentiment index). The positive sign of food inflation indicates that 92.0 per cent of food inflation in one period lagged, which explains current inflation dynamics, which is expected. In the case of non-manufacturing PMI, it is expected that increased business activities in the non-manufacturing sectors would dampen inflationary pressure. For both models, increased rainfall lagging one period has a dampening impact on food inflation as expected; rainfall affects agricultural output positively by increasing food supply during the harvest season, thereby exerting downward pressure on food prices. The sentiment index itself is correctly signed. Thus, this explains that an increase in the coefficient of the sentiment index would exert upward pressure on food inflation.

V.1.2 Core Inflation Models

Table 2 shows the coefficients of the explanatory variables, namely, the index of narrow money and non-agricultural output (M_INDEX_NARY), the exchange rate (IEW), manufacturing PMI (PMI), and the sentiment index (SENTIMENT). The one period lagged core inflation and the contemporaneous exchange rate were found to be statistically significant at the 1.0 per cent level.

Table 2: The Estimated Core Inflation Models

	Without Sentiment		**With Sentiment**	
Variable	**Coefficient**	**Prob.**	**Coefficient**	**Prob.**
CORE_INF (-1)	0.61	0.00	0.62	0.00
M_INDEX_NARY	0.00	0.91	0.00	0.95
M_INDEX_NARY (-1)	0.00	0.74	0.00	0.77
M_INDEX_NARY (-2)	0.00	0.70	0.00	0.72
M_INDEX_NARY (-3)	0.00	0.96	0.00	0.92
LOG(IEW)	12.67	0.00	12.51	0.00
LOG(PMI)	0.62	0.21	0.62	0.21
@TREND	-0.61	0.00	-0.60	0.00
@TREND^2	0.00	0.00	0.00	0.00
@SEAS (1)	0.31	0.02	0.31	0.02
SENTIMENT			0.00	0.72

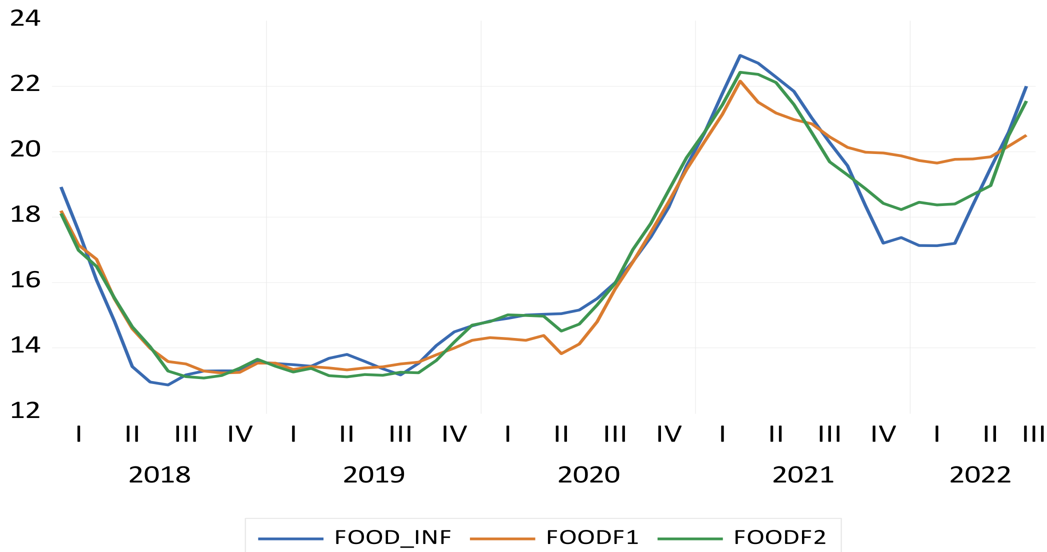
Source: Authors' estimate.

An exchange rate increase (depreciation) is expected to adversely affect core inflation by raising the price level. The parameter for (M_INDEX_NARY), explains the interaction of growth of non-agricultural production, that is, what is produced in the non-agricultural sector and narrow money growth. If this index is positive, there is expected upward pressure on core inflation. As with food inflation, the sentiment index is correctly signed. Nonetheless, an increase in the coefficient of the sentiment index would exert upward pressure on core inflation.

V.2 In-sample Forecast Evaluation

This section discusses the forecast performance of the modified STIF model with and without the computed sentiment index. An ex-post forecast analysis was conducted to ascertain the predictive power of the estimated models and examine if the inclusion of the sentiment index improves inflation forecasts for both food and core components of headline inflation. We also evaluate the accuracy of the forecasts using mean square error (MSE), which measures the quality of point forecasts by showing how close the observed data points are to the predicted values. The model was estimated for 2018M01 to 2021M12, and in-sample forecasts were generated for the same period.

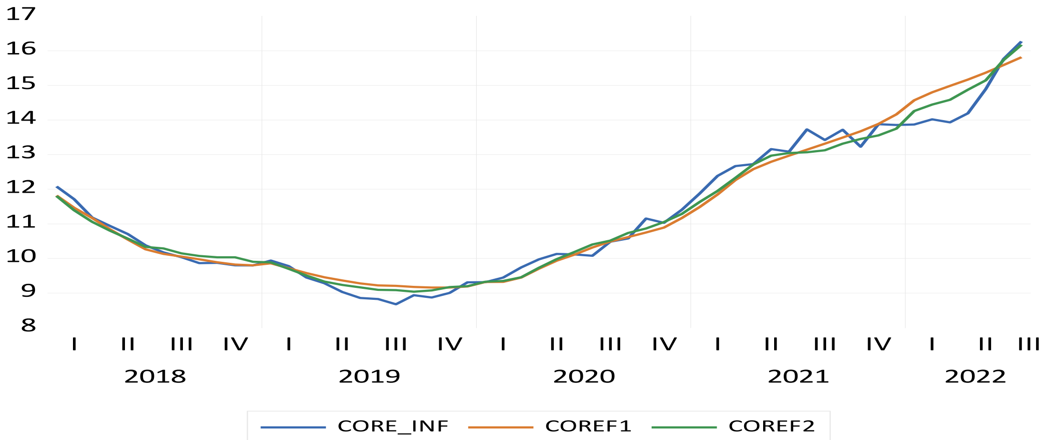
Figure 3: In-sample Forecast of Food Inflation with and without Sentiment Index



Source: Authors' estimate.

Figure 3 compares the actual inflation values with predicted values of food inflation with a sentiment index (FOODF2) and without a sentiment index (FOODF1). This reveals that the benchmark model accurately predicts food inflation and adequately captures major turning points. Likewise, the model including the Sentiment Index, accurately predicts food inflation as most turning points were successfully tracked. However, a cursory examination reveals that the model with the inclusion of the index outperforms the benchmark model. This can be seen in the peaks and troughs of the series.

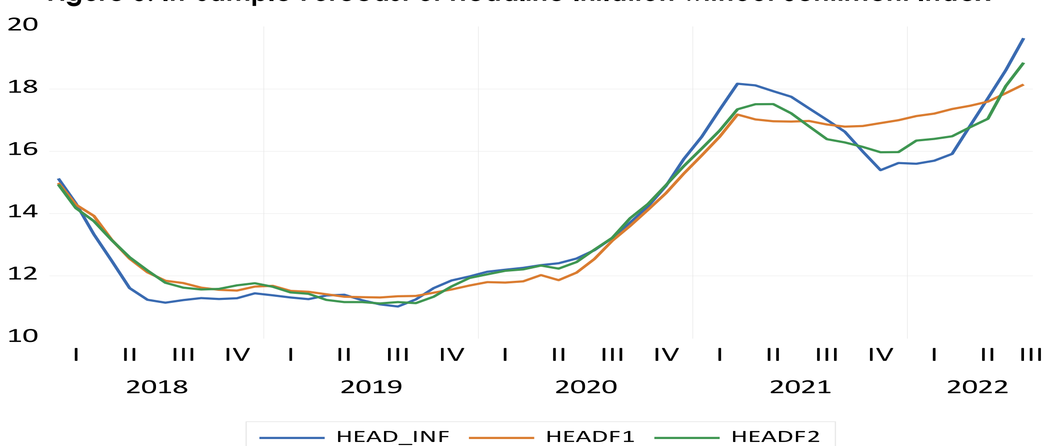
Figure 4: In-sample Forecast of Core Inflation with and without Sentiment Index



Source: Authors' estimate.

Figure 4 illustrates the forecast and actual values of core inflation with and without the Sentiment Index. Both models produce accurately predicted values for core inflation with little or no discrepancies with or without the sentiment index

Figure 5: In-sample Forecast of Headline Inflation without Sentiment Index



Source: Authors' estimate.

Figure 5 illustrates the forecast and actual values of headline inflation with and without the sentiment index. As with the predictions for both the food and core components, the headline inflation forecast closed tracked the actual values. However, a cursory examination reveals that the margin of errors is better minimised in the models with the Sentiment Index as the gaps between the actual and predicted values are narrower than that of the benchmark model. Thus, we can conclude that a Sentiment Index improves the headline inflation forecast and would enhance decision making of the Monetary Policy Committee of the Bank.

Table 3: Mean Square Errors from the In-sample Forecasts

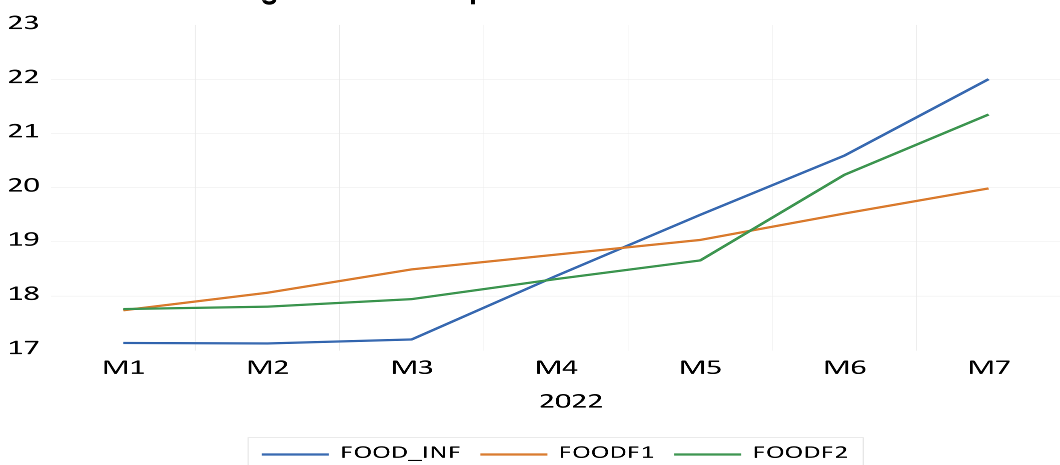
Food		Core		Headline	
Without sentiment	With sentiment	Without sentiment	With sentiment	Without sentiment	With sentiment
1.020	0.297	0.123	0.070	0.463	0.185

Source: Authors' estimate.

V.3 Out-of-sample Forecast Evaluation

In this section, we generate out-of-sample forecasts of food and core variants of inflation and compare the forecast performance under the model with and without the computed sentiments index. The forecasts are produced for the period 2022M01 to 2022M07.

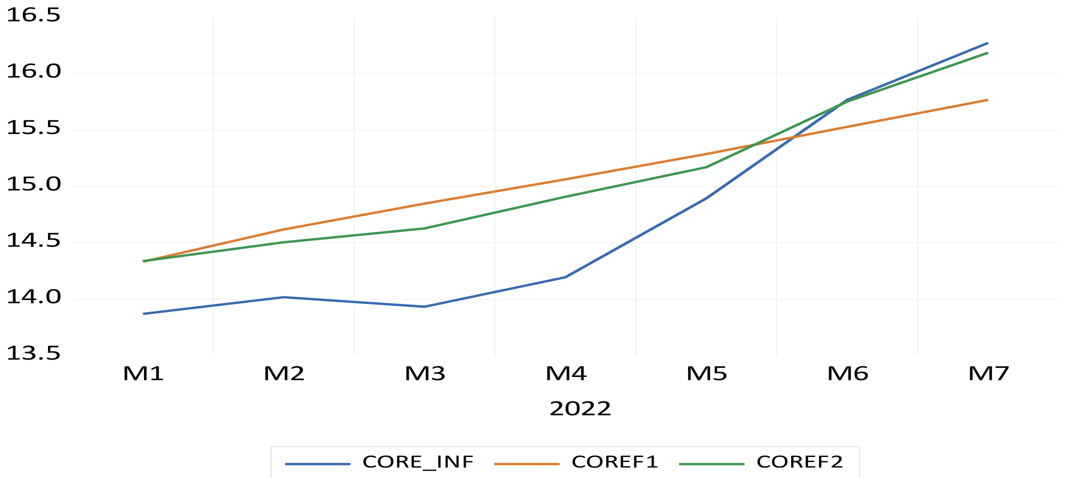
Figure 6: Out-sample Forecast of Food Inflation



Source: Authors' estimate.
Note: FOOD_INF depicts actual values of food inflation, while FOODF1 and FOODF2 represent its forecasts without and with the public sentiment variable, respectively.

Figure 6 shows the actual food inflation and its forecasts generated based on the STIF model with and without the computed sentiment index. As can be observed, the predicted values of food inflation with sentiments mimic the actual values of food inflation better than the predicted values of food inflation under the model without the sentiments variable. Overall, the usefulness of public sentiment about price developments is evident.

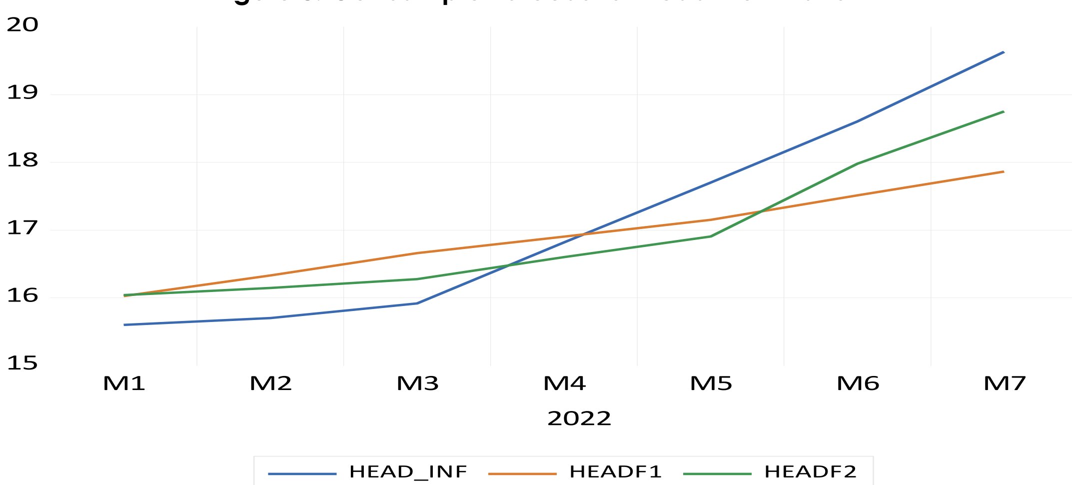
Figure 7: Out-sample Forecast of Core Inflation



Source: Authors' estimate.
Note: While CORE_INF are actual values of core inflation, COREF1 and COREF2 represent forecasts without and with sentiments, respectively.

A cursory examination also reveals that the predicted values for core inflation with sentiments closely mimic the turning points of the actual core inflation series.

Figure 8: Out-sample Forecast of Headline Inflation



Source: Authors' estimate.
Note: While HEAD_INF are actual values of headline inflation, HEADF1 and HEADF2 represent forecasts without and with sentiments, respectively.

The out-of-sample forecast of headline inflation, covering the period 2022M1 to 2022M7, is presented in Figure 8. These results further confirm the role of the Sentiment Index in predicting inflation in Nigeria.

Table 4: Mean Square Errors from the Out-sample Forecasts

Food		Core		Headline	
Without sentiment	With sentiment	Without sentiment	With sentiment	Without sentiment	With sentiment
1.21	0.38	0.38	0.22	0.82	0.34

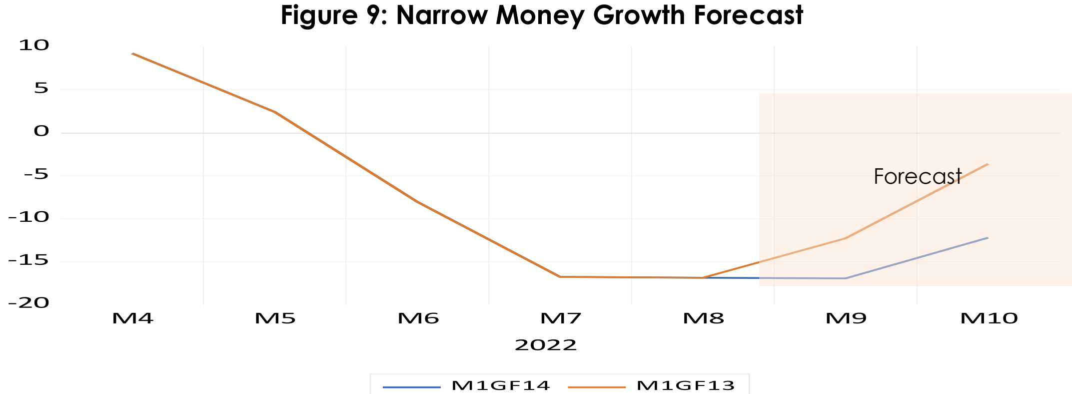
Source: Authors' estimate.

Using the RMSE criteria, it is evident that the models that include a sentiment index provides more stable and accurate forecasts than the models without the index for all components of inflation given the lower RMSE values. Of all the models, the values for the RMSE for food inflation have a larger deviation (-0.83) between the two models, which bolsters the earlier visual presentation. Thus, the sentiment index improves the forecasts of inflation in Nigeria, particularly food inflation.

V.4 Forecasting Inflation using Sentiment Index

The study already established the predictive powers of the Sentiment Index using both the in-sample and out-of-sample forecast evaluations. To conduct out-of-sample forecasts for food, core, and headline inflation for 2022M8 to 2022M10, we assume the following: MPR remains constant at 14.0 per cent across the forecast horizon, the exchange rate at the Importer's and Exporter's window remains constant at ₦417.38/US$, NPMI remains constant at 50.4, its value in July 2022, while rainfall exhibits similar cycles as previous years across the forecast horizon. Forecast was made for narrow money growth (M1G) based on a model that expressed M1G as a function of MPR, an autoregressive term, and seasonal effects. Figure 9 shows that M1G is expected to rise steadily across the forecast period, albeit with a slope if MPR remains constant at 14.0 per cent. MIGF and MIGF13 in Figure 9 represent forecast of narrow money growth when MPR was at 14.0 per cent and 13.0 per cent, respectively.

Figure 9: Narrow Money Growth Forecast



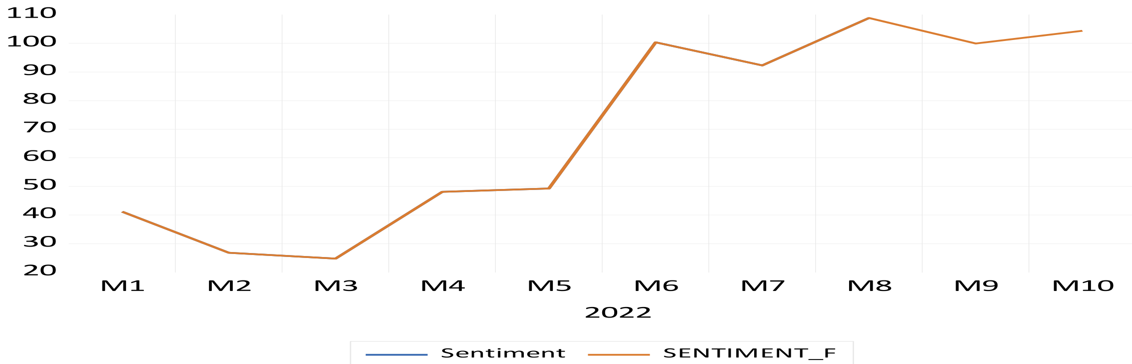
Source: Authors' estimate.

Table 5: Forecasts of Inflation

	Food Inflation (Food_INF)	Core Inflation (Core_INF)	Headline Inflation (Head_INF)
2022M8	22.91	16.74	19.81 Forecast
2022M9	23.08	17.06	20.05
2022M10	23.00	17.38	20.17

Source: Authors' estimate.

Figure 10: Sentiment Index Forecast



Source: Authors' estimate.

The forecast of the Sentiment Index reveals a moderation in its outlook across the forecast horizon (Figure 10). This implies anchoring inflation expectation through effective implementation of monetary policy will reduce public sentiment and moderate inflation outlook in the short-term.

Table 5 presents the forecasts of headline inflation and its components based on the model that includes public sentiment on price developments. The result shows that core and food inflation is expected to rise slightly in August and September. However, food inflation is expected to moderate slightly in October, while the core is expected to rise marginally in the same month. The net impact on the headline is a marginal increase in August, September, and October, 2022.

VI. Summary and Recommendation

The study applied text mining techniques to extract public sentiments about price developments in Nigeria and explored its usefulness for improving the performance of CBN's short-term inflation forecasting (STIF) model. The results showed that the computed public sentiment index correlates strongly with all the components of inflation; thus, the index provides a useful signal for future price movements in Nigeria. The results further indicated that in-sample and out-of-sample forecasts generated based on the Sentiment Index perform better than the models without the Index for all components of inflation. The study concludes that text-based indicators, such as derived from social media posts on price developments, are useful for generating better forecast of headline inflation in Nigeria. Given that consumer sentiments can potentially drive inflation expectations, the study argues that increased monetary policy communication by the CBN could help ensure that economic agents' sentiments regarding inflation are well anchored.

References

Adebiyi, M. A., Adenuga, A. O., Abeng, M. O., Omanukwe, P. N., & Ononugbo, M. C. (2010). Inflation forecasting models for Nigeria. *Central Bank of Nigeria Occasional Paper, 36*.

Adeleye, N., Ogundipe, A. A., Ogundipe, O., Ogunrinola, I., & Adediran, O. (2019). Internal and external drivers of inflation in Nigeria. *Banks Bank System*, *14*(4), 206-218.

Alessi, L., Barigozzi, M., & Capasso, M. (2009). Forecasting large datasets with conditionally heteroskedastic dynamic common factors. Working Paper No. 1115, European Central Bank.

Algaba, A., Ardia, D., Bluteau, K., Borms, S., & Boudt, K. (2020). Econometrics meets sentiment: An overview of methodology and applications. *Journal of Economic Surveys*, *34*(3), 512-547.

Ardia, D., Bluteau, K., & Boudt, K. (2019). Questioning the news about economic growth: Sparse forecasting using thousands of news-based sentiment values. *International Journal of Forecasting*, *35*(4), 1370-1386.

Askitas, N., & Zimmermann, K. F. (2009). Google econometrics and unemployment forecasting. *Applied Economics Quarterly*, *55*(2), 107–120.

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics*, *131*(4), 1593–1636.

Bańbura, M., Giannone, D., & Lenza, M. (2015). Conditional forecasts and scenario analysis with vector autoregressions for large cross-sections. *International Journal of forecasting*, *31*(3), 739-756.

Bernanke, B., Boivin, J., & Eliasz, P. S. (2005). Measuring the effects of monetary policy: A factor-augmented vector autoregressive approach. *Quart J Econ*, *120*(1),387–422.

Bordoloi, S., Biswas, D., Singh, S., Manna, U. K., & Saggar, S. (2010). Macroeconomic forecasting using dynamic factor models. *Reserve Bank of India Occasional Paper*, *31*(2), 69–83.

Box, G. E., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control*. Holden-Day.

Clements, M. P., & Reade, J. J. (2020). Forecasting and forecast narratives: The bank of England inflation reports. *International Journal of Forecasting*, *36*(4), 1488-1500.

Choi, H., & Varian, H. (2012). Predicting the present with Google trends. *Economic record*, *88*, 2-9.

Chong, k., & Tuckett, D. (2014). Constructing conviction through action and narrative: How money managers manage uncertainty and the consequences for financial market functioning. *Socio-Economic Review*, 1-26. doi:doi:10.1093/ser/mwu020.

De Caigny, A., Coussement, K., De Bock, K. W., & Lessmann, S. (2020). Incorporating textual information in customer churn prediction models based on a convolutional neural network. *International Journal of Forecasting*, *36*(4), 1563-1578.

De Mol, C., Giannone, D., & Reichlin, L. (2008). Forecasting using a large number of predictors: is Bayesian shrinkage a valid alternative to principal components? *J Econom¸146*(2),318–328.

Dominguez, K. M., & Shapiro, M. D. (2013). Forecasting the recovery from the great recession: Is this time different? *American Economic Review, 103*(3), 147-152.

Durevall, D., & Sjö, B. (2012). The dynamics of inflation in Ethiopia and Kenya. African Development Bank Group.

European Central Bank. (2017). Domestic and global drivers of inflation in the Euro area. *ECB Economic Bulletin*, (4).

Figueiredo, F. M. (2010). Forecasting Brazilian inflation using a large data set. Central Bank of Brazil Working Paper, 228.

Forbes, K. J., Hjortsoe, I., & Nenova, T. (2018) The shocks matter: Improving our estimates of exchange rate pass-through. *Journal of International Economics, 114*, 25.

Guzman, G. (2011). Internet search behaviour as an economic forecasting tool: The case of inflation expectations. *Journal of Economic and Social Measurement*, *36*(3), 119-167.

Jegadeesh, N., & Wu, D. (2013). Word power: A new approach for content analysis. *Journal of financial economics*, *110*(3), 712-729.

Jones, J. T., Sinclair, T. M., & Stekler, H. O. (2020). A textual analysis of Bank of England growth forecasts. *International Journal of Forecasting, 36*(4), 1478-1487.

Kearney, C., & Liu, S. (2014). Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis*, *33*, 171-185.

Keynes, J. M. (1936). The supply of gold. T*he Economic Journal*, *46*(183), 412-418.

Kohlscheen, E. (2021). What does machine learning say about the drivers of inflation? SSRN 3949352.

Kraaijeveld, O., & De Smedt, J. (2020). The predictive power of public Twitter sentiment for forecasting cryptocurrency prices. *Journal of International Financial Markets, Institutions and Money*, *65*, 101188.

Li, Y., Bu, H., Li, J., & Wu, J. (2020). The role of text-extracted investor sentiment in Chinese stock price prediction with the enhancement of deep learning. *International Journal of Forecasting*, *36*(4), 1541-1562.

Liang, C., Tang, L., Li, Y., & Wei, Y. (2020). Which sentiment index is more informative to forecast stock market volatility? Evidence from China. *International Review of Financial Analysis*, *71*, 101552.

Lindberg, F. (2011). Nowcasting Swedish retail sales with google search query data (Master's Thesis), Stockholm University.

McLaren, N., & Shanbhogue, R. (2011). Using internet search data as economic indicators. *Bank of England Quarterly Bulletin*, No. 2011 Q2.

Mordi, C. N., Adebiyi, M. A., & Adamgbe, E. T. (2012). Short term inflation forecasting for monetary policy in Nigeria. Central Bank of Nigeria.

Mordi, C., Essien, A., Adenuga, A. O., Omanukwe, P. N., Ononugbo, M., Oguntade, A. A., & Ajao, O. M. (2013). The dynamics of inflation in Nigeria. *A Monograph,* Central Bank of Nigeria.

Nagy, E. É., & Tengely, V. (2018). The external and domestic drivers of inflation: The case study of Hungary. *BIS Paper*, (100j).

Niesert, R. F., Oorschot, J. A., Veldhuisen, C. P., Brons, K., & Lange, R. J. (2020). Can google search data help predict macroeconomic series? *International Journal of Forecasting*, *36*(3), 1163-1172.

Nyman, R., Ormerod, P., Smuith, R., & Tuckett, D. (2014). Big data and economic forecasting: A top-down approach using directed algorithmic text analysis. Submitted to ECB Workshop on Big Data for Forecasting and Statistics. Frankfurt 7-8 April, 2014.

Pop, R. E., & Murăraşu, B. (2018). The drivers of inflation in a small open European economy: The case of Romania. *International Journal of Trade, Economics and Finance*, *9*(2).

Rambaccussing, D., & Kwiatkowski, A. (2020). Forecasting with news sentiment: Evidence with UK newspapers. *International Journal of Forecasting*, *36*(4), 1501-1516.

Ramey, V. A., & Shapiro, M. D. (1999). Costly capital reallocation and the effects of government spending. *NBER Working Papers 6283*.

Rodriguez, M. D. G., & Yoldas, E. (2016). Drivers of inflation compensation: Evidence from inflation swaps in advanced economies (No. 2016-12-30-2). Board of Governors of the Federal Reserve System (US).

Sharpe, S. A., Sinha, N. R., & Hollrah, C. A. (2022). The power of narrative sentiment in economic forecasts. *International Journal of Forecasting*. https://doi.org/10.1016/j.ijforecast.2022.04.008 0169-2070.

Simionescu, M. (2022). Econometrics of sentiments-sentometrics and machine learning: The improvement of inflation predictions in Romania using sentiment analysis. *Technological Forecasting and Social Change*, *182*, 121867.

Suchoy, T. (2009). Query indices and a 2008 downturn: Israeli data. Technical report, Bank of Israel.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of finance*, *62*(3), 1139-1168.

Tuckett, D., Ormerod, P., Nyman, R., & Smith, R. E. (2015). Improving economic prediction: A new method for measuring economic confidence and its impact on the evolution of the US economy. *Working Paper.* https://www. researchgate. net/publication/275333686.

Tule, M. K., Salisu, A. A., & Chiemeke, C. C. (2019). Can agricultural commodity prices predict Nigeria's inflation? *Journal of Commodity Markets*, 16, 100087.

Vosen, S., & Schmidt, T. (2011). Forecasting private consumption: Survey-based indicators vs. Google trends. *Journal of Forecasting*, *30*(6), 565–578.

Westerlund, J., & Narayan, P. K. (2012). Does the choice of estimator matter when forecasting returns? *Journal of Banking & Finance*, 36(9), 2632-2640.S.

Wojarnik, G. (2021). Sentiment analysis as a factor included in the forecasts of price changes in the stock exchange. *Procedia Computer Science*, 192, 3176-3183.

Xu, Q. A., Chang, V., & Jayne, C. (2022). A systematic review of social media-based sentiment analysis: Emerging trends and challenges. *Decision Analytics Journal*, 100073.