Economics Bulletin

Volume 44, Issue 3

Credit market expectations and the business cycle: evidence from a textual analysis approach

Mardoqueo Arteaga
Fordham University

Abstract

This paper examines the relationship between errors in credit spread expectations and key macroeconomic indicators over the period 1948 to 2022. By employing textual analysis on Wall Street Journal title pages, I construct a historical proxy for credit market sentiment, extending the data on credit spread expectations back to 1919. The Survey of Professional Forecasters provides the training data for this model. The analysis reveals that increases in credit spread expectation errors, interpreted as signals of heightened market optimism, are robust predictors of subsequent declines in economic activity. Most saliently, a one-standard deviation increase in forecast errors is associated with a 1.47 percentage point decline in GDP growth, highlighting the significant role of credit market sentiment in driving macroeconomic cycles.

Thanks to my dissertation committee, Fordham and Columbia University seminar participants, Roger Ferguson, and the referees for their constructive input and support. This paper uses data from the Survey of Professional Forecasters (Philadelphia Fed) and Google's word2vec model. The usual caveat on errors applies.
Citation: Mardoqueo Arteaga, (2024) "Credit market expectations and the business cycle: evidence from a textual analysis approach", *Economics Bulletin*, Volume 44, Issue 3, pages 1242-1253
Contact: Mardoqueo Arteaga - arteaga@fordham.edu.
Submitted: August 26, 2024. **Published:** September 30, 2024.

1 Introduction

Upward moves in excess bond premium – again, those corresponding to a widening of credit spreads–are very informative about the future evolution of the real economy . . . I have to believe that our macro models will ultimately be more useful as a guide to policy if they build on a more empirically realistic foundation with respect to the behavior of interest rates and credit spreads.

Governor Jeremy C. Stein, March 2014 (Stein, 2014)

The interplay between credit markets and macroeconomic cycles has long been a focal point of economic research, with credit spreads serving as key indicators of financial conditions. During periods of economic expansion, credit spreads typically narrow, reflecting reduced perceived risk, while they tend to widen during downturns, signaling increased financial strain. This cyclical behavior has led to extensive debate regarding the drivers of these fluctuations—whether they stem from fundamental economic factors or from behavioral biases within the market.

Credit spread expectations—anticipations of future changes in these spreads—are influenced by factors such as economic risk, interest rates, and investor sentiment. For instance, if investors expect higher economic risk, they may anticipate wider spreads, adjusting their strategies accordingly. Unfortunately, direct data on credit spread expectations have been limited, with key sources like the Blue Chip Financial Forecasts only beginning to collect such data in 1999.

In this paper, I provide empirical evidence on the predictive power of errors in credit spread expectations—differences between realized credit spreads and their forecasts—on macroeconomic outcomes. By deriving my own proxy for credit spread expectations through a textual analysis of Wall Street Journal front pages from 1919 onwards, I construct a historical series of these expectations. This data enables an empirical investigation into the relationship between expectation errors and macroeconomic indicators such as GDP, unemployment, and private domestic investment from 1948 to 2022.

I focus on the BAA credit spread, defined as the difference between the yield on BAA-rated corporate bonds and the risk-free 10-Year Treasury Bond Yield. Credit spread dynamics offer critical insights into investor perceptions of risk and the broader economic outlook. Wide spreads indicate that investors demand higher returns to compensate for perceived increased risk, often signaling economic uncertainty. Conversely, narrow spreads reflect confidence and willingness to accept lower returns, typically associated with economic stability.

The findings indicate that periods of elevated expectation errors, which reflect overly optimistic market sentiment, are followed by significant shifts in macroeconomic indicators. Specifically, a one-standard deviation increase in these errors predicts a 1.47 percentage point decline in GDP growth over the subsequent year, a 0.95 percentage point increase in the unemployment rate, and a 1.42 percentage point reduction in private domestic investment growth. These effects, which vary in timing across different indicators, suggest that labor markets respond more immediately to deteriorating credit conditions, while GDP growth and investment decisions exhibit a delayed response. The results are robust to various controls and support the behavioral view that sentiment-driven fluctuations in credit markets substantially influence broader economic cycles.

The next section situates this study within the broader research context. Section 3 details the data sources and the construction of the explanatory variables. Section 4 describes the empirical strategy. The results are presented in Section 5, and Section 6 concludes with a discussion of the implications of these findings.

2 Behavioral and Textual Analysis in Credit Markets

This paper contributes to two active areas of research. The first is the behavioral approach to understanding business cycles through credit spreads. Recent literature, such as Bordalo et al. (2018), highlights how credit spreads can reflect investor sentiment rather than purely rational expectations. Using data from the Blue Chip Financial Forecasts, Bordalo et al. demonstrate that credit spread expectation errors—differences between predicted and actual spreads—are often predictable and linked to economic cycles. Their findings suggest that when credit spreads are low, forecasts tend to be overly optimistic, leading to larger errors and subsequent economic downturns.

My analysis applies similar methods to data from the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters (SPF) from 2010 to 2022. Figure 1 shows that forecast errors (gray line) are predictable: low credit spreads (black line) lead to positive errors (overly optimistic forecasts), and high spreads result in negative errors (overly pessimistic forecasts), reflecting sentiment-driven cycles in the market. Table 1 formalizes this through predictive regressions, showing a relationship between the current credit spread and the actual spread (Column 1), forecasted spread (Column 2), and forecast error (Column 3). The significant coefficients across all models confirms that forecast biases are linked to current spreads: higher spreads lead to pessimistic forecasts, and lower spreads to optimistic forecasts.

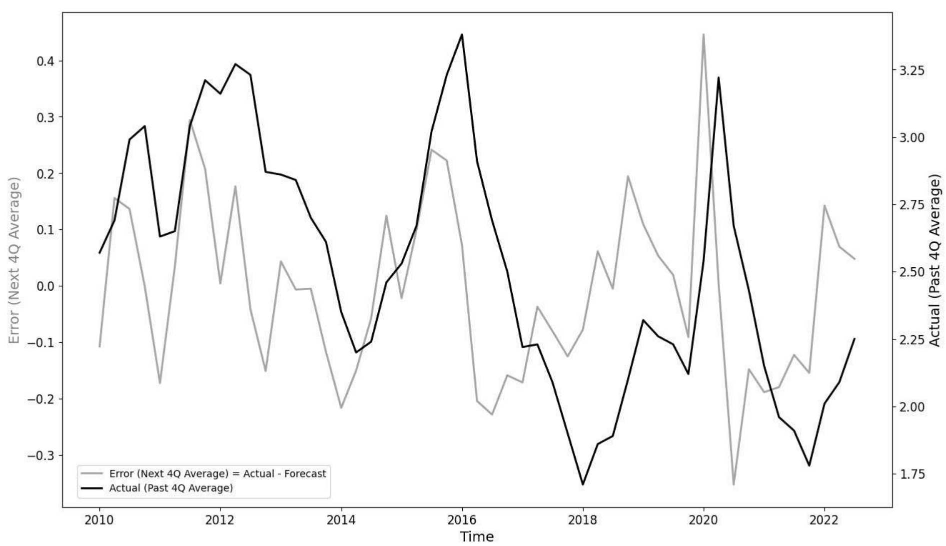

Figure 1: Quarterly time series plot adapted from Figure 1 in Bordalo et al. (2018). The gray line represents the average credit spread expectation errors (actual minus forecast) over quarters $t+1$ to $t+4$ (left scale), while the black line shows the average actual credit spread over quarters $t-4$ to $t-1$, where $t-1$ is the most recent quarter prior to the forecast (right scale). Survey of Professional Forecasters, 2010Q1 to 2022Q3.

Table 1. Predictability Tests on Credit Spreads (Actual, Forecast, and Error)

	Actual Spread (1)	Forecast Spread (2)	Error (3)
Current Spread	0.32*	0.51***	-0.31***
	(0.19)	(0.11)	(0.08)
Constant	1.16**	1.23**	1.82***
	(0.49)	(0.51)	(0.37)
Total Observations	51	51	51
R^2	0.64	0.54	0.26

Notes: Quarterly regressions following Table I in Bordalo et al. (2018). The independent variable is the actual credit spread averaged over quarters $t-4$ to $t-1$. (1) is the actual spread averaged between $t+1$ and $t+4$, (2) is the forecasted spread averaged between $t+1$ and $t+4$, and (3) is the forecast error (actual minus forecasted spread). Data are from the Survey of Professional Forecasters (2010Q1–2022Q3). Newey-West standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

This study closely aligns with the work of Gilchrist and Zakrajsek (2012) and López-Salido et al. (2017). Gilchrist and Zakrajsek show that credit spreads contain valuable information about economic conditions, with variations in risk premiums being linked to investor sentiment. López-Salido et al. focus on how elevated sentiment, as measured by the expected return to bearing credit risk, predicts economic downturns. My approach combines the construction of a credit spread expectation proxy using textual analysis, as pioneered by Gilchrist and Zakrajsek, with the broader time span and predictive framework used by López-Salido et al.. By constructing a historical series of credit spread expectations through textual analysis of Wall Street Journal front pages, I provide a direct measure of market sentiment that can be used to predict macroeconomic outcomes.

The second research strand involves applying text analysis and machine learning to economics and finance. As detailed by Gentzkow et al. (2019), this approach involves representing text as numerical arrays, mapping these arrays to predicted outcomes, and using these predictions in both descriptive and causal analyses. Applications have ranged from extracting sentiment from financial statements (Ke et al., 2019; Yue and Jing, 2022) to analyzing central bank communications (Hansen et al., 2018; Shapiro and Wilson, 2019; Doh et al., 2020). My contribution to this literature involves using vector embedding and cosine similarity (Doh et al., 2020) to analyze Wall Street Journal front pages, employing Locality Sensitive Hashing (LSH) for clustering (Andoni et al., 2015), and using Latent Dirichlet Allocation (LDA) to uncover latent themes in the text (Hansen et al., 2018). This methodology allows me to create a robust proxy for credit spread sentiment, which I use to predict macroeconomic indicators.

3 Data and Construction of Explanatory Variables

This section details the data sources and the construction of the explanatory variables, focusing on how textual factors derived from Wall Street Journal (WSJ) front pages serve as proxies for credit market sentiment. These factors are crucial for predicting historical credit spread expectations and their errors.

3.1 Acquired Data

The primary data used in this study include realized BAA credit spreads from January 1919 to September 2022, obtained from the Federal Reserve Economic Database (FRED) and Capital Markets Data. The BAA credit spread, defined as the difference between Moody's seasoned BAA Corporate Bond Yield and the 10-Year Treasury Bond Yield, serves as a key indicator of credit market conditions. This measure is calculated from data available since January 1919 and continues through the most recent complete quarter of 2022.

In addition, I utilize credit spread expectations from the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters (SPF), which spans from January 2010 to September 2022. The SPF provides quarterly forecasts of the BAA Corporate Bond Yield for up to four quarters ahead. These forecasts, denoted as $CS_{t+\tau|t-1}$ for $\tau = 0, 1, 2, 3, 4$, allow me to calculate the BAA credit spread expectations by pairing them with the realized 10-Year Treasury Bond Yield.

For the analysis of macroeconomic indicators such as GDP, unemployment, CPI inflation, and Treasury yields, I rely on data sourced from FRED and ALFRED, covering the period from 1948Q1 to 2022Q3.

3.2 Constructed Data — *The Wall Street Journal*

To extend the historical series of credit spread expectations, I perform textual analysis on WSJ front pages from January 1919 to September 2022. By utilizing Amazon Web Services' Textract for Optical Character Recognition (OCR), I extract text from WSJ images to generate textual factors, which serve as proxies for historical credit spread sentiment.

The construction of these textual factors follows a three-step process: word embedding, clustering, and topic modeling. These steps are designed to reduce the dimensionality of the text data and create a set of textual factors that quantitatively represent sentiment over time.

3.2.1 Word Embedding

The initial stage in textual analysis involves summarizing or representing the words present in the texts through word embedding. Traditional methods, such as one-hot vector encoding, treat words as independent units, which can lead to inaccuracies in capturing their semantic relationships. To overcome this, I employ the Word2Vec model, which generates real-valued vector representations for words based on their context.

Specifically, I use the Word2Vec model pre-trained by Google on its Google News dataset. This model creates a learned embedding vector $w \in \mathbb{R}^{p \times V}$, where p is the dimension of the vector (set to 300 in this analysis), and V is the size of the document vocabulary. The semantic similarity between any two word vectors, w_i and w_j, is measured using cosine similarity, defined as:

$$\text{similarity}(w_i, w_j) = \arccos\frac{\langle w_i, w_j\rangle}{|| w_i || || w_j ||} = \frac{\sum_{i=1}^{n} w_i \sum_{j=1}^{n} w_j}{\sqrt{\sum_{i=1}^{n} w_i^2}\sqrt{\sum_{j=1}^{n} w_j^2}} \tag{1}$$

where $\langle w_i, w_j\rangle$ represents the dot product of the two vectors, and $|| w_i ||$ and $|| w_j ||$ denote their magnitudes. This similarity score ranges from -1 (indicating total dissimilarity) to 1 (indicating identical meaning), enabling the model to capture subtle semantic relationships between words based on their contextual usage.

3.2.2 Clustering

To manage the high dimensionality of the word vectors generated in the previous step, I apply Locality Sensitive Hashing (LSH) to cluster similar words together. LSH is a technique that categorizes data into bins based on their proximity in vector space, effectively grouping similar words.

The clustering process is governed by a family of hash functions H. For any two vectors w_i and w_j, the probability that they are placed in the same bin by a random hash function $h(\cdot) \in H$ is given by:

$$P[h(w_i) = h(w_j)] = p_1, \quad \text{for any } w_i, w_j \text{ such that } d(w_i, w_j) \leq d_1 \tag{2}$$

$$P[h(w_i) = h(w_j)] = p_2, \quad \text{for any } w_i, w_j \text{ such that } d(w_i, w_j) \geq d_2 \tag{3}$$

where $d(\cdot)$ is the distance metric (cosine similarity in this case), and p_1 and p_2 represent the probabilities of the vectors being similar or dissimilar, respectively. The goal is to maximize p_1 (minimizing false negatives) and minimize p_2 (minimizing false positives). To implement this, I use the random hyperplane projection method, where the hash function for a vector w is defined as:

$$h_r(w) = sgn(\langle r, w\rangle), \quad \text{for } r \text{ randomly sampled from the unit sphere } S^{p-1} \tag{4}$$

This method generates a spherically symmetric random vector r of unit length, which is used to cluster the high-dimensional word embeddings into a manageable number of bins or clusters K.

3.2.3 Topic Modeling

The final step involves reducing the dimensionality of the clustered word vectors by employing Latent Dirichlet Allocation (LDA) for topic modeling. LDA is a statistical technique that identifies the underlying semantic structure of a text by modeling the distribution of words across a set of topics.

In this analysis, I assume a simple, two-distribution data-generating process. Each WSJ front page is represented by a distribution over a collection of topics, while each topic is represented by a distribution of words. Mathematically, this is expressed as:

$$[\Theta B]_{dw} := P(w_{di} = w \mid [\theta_d, \beta_1, \beta_2, \ldots, \beta_K]) = \sum_k \theta_{dk}\beta_{kw} \tag{5}$$

where $\Theta = [\theta_1, \theta_2, \ldots, \theta_D]' \in \mathbb{R}^{D\times K}$ represents the probability that document d covers a certain topic, $B = [\beta_1, \beta_2, \ldots, \beta_K]' \in \mathbb{R}^{K\times V}$ represents the probability distribution over words for each topic, and N_{dw} denotes the number of times word w appears in document d. The parameters $\theta_d \sim \text{Dirichlet}(\alpha)$ and $\beta_k \sim \text{Dirichlet}(\eta)$ are sampled from Dirichlet distributions, with α and η being hyperparameters that control the sparsity of the distributions.

This approach enables me to generate a set of textual factors that represent the primary themes within the WSJ text data, which can be used to estimate credit spread expectations and their errors over time.

3.2.4 Validation of Factors

The textual factors produced through the above steps are validated by calculating their loadings, which quantify the degree to which each WSJ front page aligns with a particular topic. The loading x_i^d for document d on factor i is computed using the projection:

$$x_i^d := \frac{\langle N_{S_i}^d, F_i \rangle}{\langle F_i, F_i \rangle} \tag{6}$$

where $N_{S_i}^d$ represents the word support for the factor i, and F_i is the vector representation of the factor. These loadings provide a quantitative measure of how much each document contributes to the identified topics, which are then used in regression analyses to estimate historical credit spread expectation errors.

This comprehensive approach allows for the reconstruction of a historical series of credit spread expectations and their errors, enabling a deeper analysis of their relationship with macroeconomic indicators. To grasp the concept of textual factor loadings, consider a comparison with structured data. While structured data provide quantifiable metrics, such as stock prices or the 10-Year Treasury Bond Yield, unstructured data rely on specific vocabulary to convey sentiment or themes. For instance, during the 2008 Financial Crisis, the Wall Street Journal might have frequently used terms like *financial crisis*, *investor confidence*, *fiscal policy*, and *bankruptcy*. The textual factor loadings x_i^d in Equation 6 quantitatively measure how much each document d reflects a particular topic, represented by loadings $x_1^d, x_2^d, \ldots x_K^d \in \mathbb{R}^K$.

4 Methodology for Credit Spreads and Macroeconomic Indicators

This methodology for this study involves two steps: predicting historical BAA credit spreads and assessing their impact on macroeconomic indicators. First, I apply Bordalo et al. (2018)'s approach, enhanced with textual factors from Equation 6, to forecast spreads from 1919Q1 to 2009Q4 using SPF data from 2010Q1 to 2022Q3 (Equation 7). The resulting credit spread expectation errors are then used as the independent variable (Equation 8) to evaluate their influence on macroeconomic outcomes over different time horizons.

4.1 Historical Credit Spread Expectations

To reconstruct the historical BAA credit spread expectations from 1919Q1 to 2009Q4, I first train a model using Survey of Professional Forecasters (SPF) data from 2010Q1 to 2022Q3 and enhanced by incorporating textual factor loadings derived from the Wall Street Journal front pages.

The credit spread expectations $E[CS]_t$ for quarter t are estimated using a penalized Lasso regression, which helps identify the most significant textual factors while minimizing overfitting. The regression is specified as:

$$E[CS]_t = \alpha + \mathbf{\Gamma}\mathbf{x}_t^D + \eta_t, \quad t = 1, 2, \ldots, T \tag{7}$$

Here, $\mathbf{x}_t^D$ represents the vector of textual factor loadings, $\mathbf{\Gamma}$ is a vector of regression coefficients associated with these factors, and $T = 51$ is the number of observations in the SPF sample. This approach is compared against the simpler model of Bordalo et al. (2018), where $E[CS]_t = \alpha + \gamma[CS]_t + \epsilon_t$.

Once the model is trained on the SPF data, it is applied to the entire span of the textual factor loadings from 1919Q1 to 2009Q4, producing the predicted credit spread expectations. The credit spread expectation error $\widehat{E[CSE]}_t$ is then calculated by subtracting the realized BAA credit spread from the predicted spread.

4.2 Predicting Macroeconomic Indicators

With the historical credit spread expectation errors $\widehat{E[CSE]}_t$ in hand, the next step is to assess their predictive power for macroeconomic indicators such as GDP growth, unemployment, and private domestic investment.

Following the methodology of López-Salido et al. (2017), I estimate the relationship between changes in credit spread expectation errors and these macroeconomic outcomes using the following predictive regression:

$$\Delta y_{t+h} = \beta_0 + \beta_1 \Delta \widehat{E[CSE]}_t + \boldsymbol{\gamma}'\mathbf{x_{t-1}} + \upsilon_t \tag{8}$$

In this equation, Δy_{t+h} represents the change in the macroeconomic indicator over a horizon h, where $h = 0, 1, 2, 3, 4$ quarters. The independent variable $\Delta \widehat{E[CSE]}_t$ is the change in the credit spread expectation error from quarter $t-1$ to t. The control variables $\mathbf{x_{t-1}}$ include the credit spread in quarter $t-1$, lagged values of the dependent variable Δy_t, CPI inflation, and changes in the 3-month and 10-year Treasury yields. The vector $\boldsymbol{\gamma}'$ represents the coefficients associated with these control variables. Newey-West standard errors are applied to correct for heteroskedasticity and autocorrelation, using the automatic lag-selection procedure of Newey and West (1994).

This approach is primarily predictive, offering insights into potential future paths of macroeconomic indicators based on the observed changes in credit spread expectation errors. For instance, a negative and statistically significant coefficient β_1 suggests that an increase in the credit spread expectation error predicts a decline in GDP growth.

A key concern in the estimation of Equation 8 is the possibility of omitted variable bias (OVB). Specifically, one might argue that the expected error of real GDP growth could be a relevant omitted variable, influencing the real GDP growth rate in period $t + h$ while also being correlated with the credit spread expectation error. If this were the case, the estimated impact of credit spread expectation errors on GDP growth might be confounded by the omitted GDP growth expectations.

To address this, I perform robustness checks by controlling for GDP growth expectations, where available. Additionally, the absence of direct data on historical GDP expectations over the full sample limits the ability to explicitly test for OVB in earlier periods. However, the inclusion of control variables for lagged GDP growth, unemployment, and Treasury yields in Equation 8 partially mitigates this concern, as these variables are highly correlated with real economic expectations.

To provide a clearer interpretation of the regression coefficients, Table 2 summarizes the descriptive statistics for key macroeconomic variables used in the analysis. This includes the actual credit spread expectation errors (Actual Error) and the predicted credit spread expectation errors calculated by subtracting the realized BAA credit spread from the predicted spread from Equation 7 (Predicted Error), along with the growth rates for GDP and Domestic Investment (GPDI), and the unemployment rate.

Table 2. Descriptive Statistics of Key Macroeconomic Variables

	Actual Error	Predicted Error	GDP Growth	GPDI Growth	Unemployment Rate
Average	0.355	0.183	3.169	4.687	5.726
St. Dev.	0.511	0.506	2.662	11.722	1.705
Obs.	51	415	299	299	299

Notes: Actual credit spread errors from 2010Q1-2022Q2 using the SPF. Predicted credit spread errors using textual factors from 1919Q1-2022Q2. GDP Growth, GPDI Growth, and the Unemployment Rate from 1948Q1-2022Q3.

5 Results

Figure 2 illustrates the time series of actual and predicted expectation errors. To assess the effectiveness of the textual factors in predicting BAA credit spread expectation errors, I compare the results against the benchmark model from Bordalo et al. (2018), where expectation errors are predicted based on current credit spreads. When incorporating textual factors, the current credit spread remains a significant predictor of errors, with a coefficient of 0.456 (standard error of 0.079). Without textual factors, the coefficient is slightly higher at 0.564 (standard error of 0.093). This comparison suggests that textual factors modestly enhance the model's predictive power.

To evaluate the model's performance, I conducted K-fold cross-validation, splitting the sample into 10 groups. The in-sample R^2 increases from 0.57 without textual factors to 0.69 with them, a 21.5% improvement. Out-of-sample R^2, which assesses the model's predictive accuracy on unseen data, increases from 0.41 to 0.52 with textual factors, marking a 26.8% improvement.

Table 3 presents the results of predictive regressions for three key macroeconomic indicators: real GDP growth, unemployment, and private domestic investment growth. The table focuses solely on statistically significant results to maintain clarity.

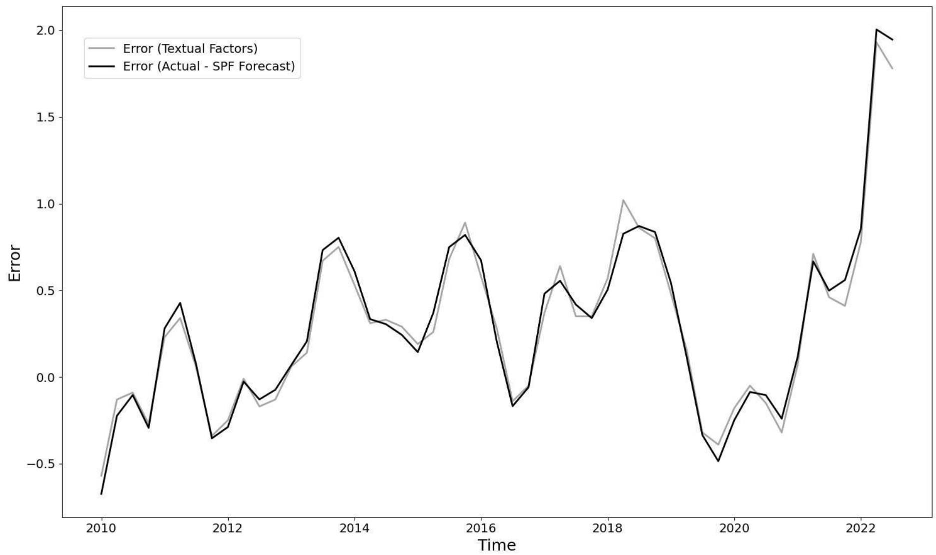

Figure 2: Quarterly time series plot, 2010Q1 - 2022Q3. The black line shows credit spread expectation errors (actual minus forecast) from consensus forecasts computed from the Survey of Professional Forecasters. The gray line shows the predicted credit spread expectation errors calculated using the textual factors.

At the four-quarter horizon ($h = 4$), a one-unit increase in the predicted expectation error of credit spreads is associated with a 2.9 percentage point decline in GDP growth. Considering the standard deviation of the expectation error (0.506), a one-standard deviation increase in the credit spread expectation error predicts a 1.47 percentage point decline in GDP growth over a one-year period. This result aligns with previous findings by Bordalo et al. (2018), supporting the notion that market sentiment biases have measurable effects on economic outcomes. The positive coefficient for the current credit spread indicates that more favorable credit conditions tend to support GDP growth.

For unemployment, the credit spread expectation error exhibits statistically significant effects across all horizons, from $h = 0$ (immediate impact) to $h = 4$ (four quarters ahead). A one-unit increase in the expectation error predicts an increase in the unemployment rate by 1.68 percentage points on average over the next year (an average calculated using the coefficients across the nowcast and subsequent quarters). Given that the standard deviation of the error, a one-standard deviation increase in the credit spread error corresponds to approximately a 0.84 percentage point increase in the unemployment rate over the next year. Given the standard deviation of the unemployment rate itself (1.705), this represents a substantial response, indicating that labor markets react quickly and significantly to deteriorating credit conditions. The negative coefficient for the current credit spread suggests that tighter credit conditions (narrower spreads) are associated with reductions in unemployment at the four-quarter horizon.

Last, at the four-quarter horizon, a one-unit increase in the credit spread expectation error predicts a 2.8 percentage point decline in private domestic investment growth. A one-standard deviation increase in the expectation error would lead to a 1.42 percentage point decline in investment growth, highlighting the role of sentiment-driven biases in shaping investment decisions. The current credit spread negatively impacts investment growth across multiple horizons, underscoring the long-term importance of stable credit conditions for sustaining capital expenditures.

Table 3. Predictive Regressions for Macroeconomic Indicators, h quarters ahead

Real GDP Growth	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
Δ Predicted Expectation Error					-0.029***
					(0.010)
Most Recent Credit Spread					0.014**
					(0.006)
Most Recent log Δ GDP	0.459***	0.516***	0.613***	0.514***	0.556***
	(0.084)	(0.080)	(0.091)	(0.063)	(0.079)
Change in Unemployment	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
Δ Predicted Expectation Error	1.871**	1.654**	1.665**	1.598**	1.601**
	(0.903)	(0.815)	(0.820)	(0.781)	(0.766)
Most Recent Credit Spread					-0.509**
					(0.219)
Most Recent Δ Unemployment	0.761***	0.707***	0.699***	0.774***	0.797***
	(0.292)	(0.264)	(0.248)	(0.282)	(0.287)
Domestic Investment Growth	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$
Δ Predicted Expectation Error					-0.028***
					(0.011)
Most Recent Credit Spread	-0.014**	-0.016**	-0.011**	-0.018**	-0.020**
	(0.007)	(0.007)	(0.005)	(0.008)	(0.010)
Most Recent log Δ Investment	0.185**	0.197*	0.251*	0.213*	0.194*
	(0.082)	(0.116)	(0.137)	(0.112)	(0.104)

Notes: Regressions cover 1948Q1–2022Q3, with a constant included but not reported. Variables are Δ Predicted Expectation Error (change from $t-1$ to t), Most Recent Credit Spread ($t-1$), Most Recent log Δ (GDP, Unemployment, or Investment from $t-2$ to $t-1$), Most Recent CPI ($t-1$), Δ 3-Month and Δ 10-Year Treasury Yields. Newey-West standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

One of the key findings in this study is the difference in the timing of how credit spread expectation errors impact various macroeconomic indicators. The results show an immediate impact on the unemployment rate but a lagged effect on GDP growth and private investment. This pattern can be explained by the different transmission channels through which credit conditions affect economic activity.

Unemployment tends to respond quickly to changes in credit spreads because firms facing increased borrowing costs or restricted access to credit adjust labor costs almost immediately by cutting hours or laying off workers. As labor is a more flexible and variable cost relative to investment, firms often prioritize labor adjustments in response to short-term financial constraints or expected economic downturns. This could explain why the expectation error in credit spreads has an immediate impact on unemployment.

In contrast, GDP growth and private domestic investment typically exhibit lagged responses to changes in credit conditions. Investment decisions, which constitute a significant portion of GDP, involve long-term commitments of capital and are less responsive to short-term fluctuations in credit spreads. Firms may delay or scale back investment projects in response to worsening credit conditions, but such decisions often require time due to planning, financing, and project execution stages. Similarly, the broader economy (as captured by GDP growth) reacts more slowly because it aggregates a wide range of economic activities, many of which are not as immediately responsive to changes in credit conditions as labor markets.

In this sense, the immediate effect on unemployment followed by lagged effects on GDP and investment aligns with the typical sequence of adjustments observed during economic downturns, where firms first manage labor costs before altering capital expenditure plans or facing broader declines in economic output.

6 Final Remarks

This study investigates the predictive power of credit spread expectation errors on macroeconomic indicators, focusing specifically on the BAA credit spread as a proxy for investor sentiment and economic health. The analysis highlights two prevailing approaches in the literature: one that addresses the amplification of shocks through fundamental factors and another that examines the role of non-rational beliefs, such as excessive optimism, in driving business cycles.

The analysis employs data from the Survey of Professional Forecasters (2010-2022) to motivate the systematic biases narrative. By training a machine learning model with textual analysis of Wall Street Journal title pages, the study generates textual factors that serve as proxies for sentiment. These factors enable the construction of a historical series of credit spread expectation errors, which are used to predict key macroeconomic indicators.

The findings suggest that increases in credit spread expectation errors—reflecting over-optimism in the credit market—predict downturns in economic activity, including declines in GDP growth, increases in unemployment, and reductions in private domestic investment. These results underline the importance of integrating credit market sentiment into forecasting and policy-making. However, the analysis has limitations. The sentiment proxies derived from textual analysis provide insights into market behavior but do not fully capture the underlying drivers or transmission mechanisms. Future research should explore these channels in more detail to improve the precision of policy recommendations.

Overall, the results underscore the need for continued development of analytical frameworks that account for market sentiment and its impact on macroeconomic outcomes, particularly in the context of ongoing economic uncertainty.

References

Andoni, A., Indyk, P., Laarhoven, T., Razenshteyn, I., and Schmidt, L. (2015). Practical and optimal lsh for angular distance. In Cortes, C., Lawrence, N., Lee, D., Sugiyama, M., and Garnett, R., editors, *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2018). Diagnostic Expectations and Credit Cycles. *Journal of Finance*, 73(1):199–227.

Doh, T., Song, D., and Yang, S.-K. X. (2020). Deciphering Federal Reserve Communication via Text Analysis of Alternative FOMC Statements. Research Working Paper RWP 20-14, Federal Reserve Bank of Kansas City.

Gentzkow, M., Kelly, B., and Taddy, M. (2019). Text as Data. *Journal of Economic Literature*, 57(3):535–574.

Gilchrist, S. and Zakrajsek, E. (2012). Credit Spreads and Business Cycle Fluctuations. *American Economic Review*, 102(4):1692–1720.

Hansen, S., McMahon, M., and Prat, A. (2018). Transparency and Deliberation Within the FOMC: A Computational Linguistics Approach. *The Quarterly Journal of Economics*, 133(2):801–870.

Ke, Z. T., Kelly, B. T., and Xiu, D. (2019). Predicting Returns With Text Data. NBER Working Papers 26186, National Bureau of Economic Research, Inc.

López-Salido, D., Stein, J. C., and Zakrajšek, E. (2017). Credit-Market Sentiment and the Business Cycle. *The Quarterly Journal of Economics*, 132(3):1373–1426.

Newey, W. K. and West, K. D. (1994). Automatic Lag Selection in Covariance Matrix Estimation. *Review of Economic Studies*, 61(4):631–653.

Shapiro, A. H. and Wilson, D. J. (2019). The Evolution of the FOMC's Explicit Inflation Target. *FRBSF Economic Letter*.

Stein, J. C. (2014). Incorporating Financial Stability Considerations into a Monetary Policy Framework : a speech at the International Research Forum on Monetary Policy, Washington, D.C., March 21, 2014. Speech 796, Board of Governors of the Federal Reserve System (U.S.).

Yue, T. and Jing, M. (2022). Sentiment Analysis Based on Bert and Transformer. In Yang, H., Qiu, R., and Chen, W., editors, *AI and Analytics for Public Health*, Springer Proceedings in Business and Economics, pages 317–328. Springer.