Do Social Media Discussions about Government Spending Affect Consumer Confidence?

Egor Dolgov

1 Introduction

Consumer confidence is a crucial indicator for governments when making policy decisions, given that stronger confidence often leads to increased household spending and, in turn, higher economic growth (Fisher, 1993; Ludvigson, 2004). Traditionally, surveys and indices such as the Consumer Confidence Index have been used to capture how households feel about the economy. However, the rise of social media in the past decade presents an alternative lens, offering near-instant reflections of public mood on political, social, and economic issues (Bollen, Mao, & Zeng, 2011; O'Connor et al., 2010). Recent data from the Federal Reserve and the Bureau of Economic Analysis further underscore the importance of these real-time sentiment indicators during the COVID-19 pandemic, when government relief packages and stimulus checks became hotly debated topics (Smith, Brown, & Taylor, 2022). Against this backdrop, the economic question guiding this paper asks how social media discussions of government spending interact with shifts in consumer confidence, particularly during the 2020 U.S. elections. Previous research shows that text-based sentiment can predict both financial market movements and political outcomes. Scholars have linked media tone and investor attention to stock market behavior (Tetlock, 2007; Da, Engelberg, & Gao, 2011) and demonstrated how investor sentiment shapes asset prices (Barberis, Shleifer, Wurgler, 2005). As attention turned to social media, studies revealed that online sentiment often foreshadows economic activity (Garcia, 2013; Preis, Moat, & Stanley, 2013), with Twitter engagement correlating closely to broader economic and political developments (Zhang, Fuehres, & Gloor, 2011; Oh, Park, & Rao, 2015). Researchers have also begun to look at how political messaging shapes economic perceptions, noting that partisan framing of fiscal policy (Osterloh, 2018; Öztürk, 2024) and social media engagement metrics (Kim, Mikalef, & Pappas, 2014) may influence consumer behavior (Smales, 2020; Gao & Jiang, 2017). Despite these advances, we still lack a clear understanding of how social media discussions specifically focused on government spending align with changing consumer sentiment. This paper contributes to the literature by examining 2020 election-related tweets in the

United States, focusing on the volume and sentiment of spending-related discourse as possible indicators of consumer confidence. By integrating large-scale social media data with established economic measures, we offer new insights into whether digital political conversations mirror or even drive real-time shifts in households' perceptions. In what follows, we outline our data sources and methods, present the core findings, and discuss their implications for policymakers, market participants, and researchers interested in the evolving relationship between online discourse and economic perceptions.

2 Data

Our analysis is based on two complementary data sources. The first source is a comprehensive dataset of Twitter posts collected during the 2020 U.S. elections. We focused on tweets that explicitly mention fiscal policy, government spending, and related keywords (e.g., "stimulus," "relief package," "spending"). These tweets were processed using sentiment analysis algorithms to generate sentiment scores, and engagement metrics such as likes, retweets, and replies were recorded. The observations are aggregated at the regional level—based on geotagging or inferred location data—allowing us to capture local variations in online discourse.

The supplemental dataset I chose is "University of Michigan Consumer Sentiment Index", which displays the Consumer Sentiment Index (CSI) from 1954 to 2024, where the values range from 0 to 100 respectively. Data from 2020 and 2021 will only be used. The assumption for the project is that CSI is a valid metric to assess consumer confidence.

Another supplemental dataset is "Household Income by State". This dataset displays the median income for each state in the United States in 2020. It will be used to check for regional economic factors that might influence consumer sentiment and confidence. Since there were only 50 data points, I transformed it into a dictionary of entries to ease the analysis.

By merging these datasets, we create a panel that links digital sentiment with conventional economic outcomes. The level of observation at the regional or state level enables us to explore both cross-sectional differences and trends over the election period. Special attention is given to controlling for standard macroeconomic variables such as unemployment rates, average income levels, and regional demographic characteristics to isolate the impact of social media sentiment on consumer confidence.

3 Summary Statistics

Summary statistics reveal a positive skew in the distribution of sentiment scores, suggesting that, on average, Twitter users expressed optimism about government spending initiatives. This is in line with the hypothesis that digital conversations can serve as a proxy for broader economic sentiment. In our subsequent regression analysis, these visual insights are reinforced by statistically significant relationships between our independent variables (tweet volume and sentiment) and the dependent variable (consumer confidence).

The following variables were chosen for the project:

- y – **Consumer Sentiment Index (CSI):** the consumer confidence index that we would correlate to the changes in other variables.
- x_1 – **Tweet Volume:** the number of tweets per region mentioning government spending-related terms. This serves as a proxy for the intensity of public discussion.
- x_2 – **Average Likes:** the mean number of likes per tweet, indicating how positively users engage with spending-related content.
- x_3 – **Average Retweets:** the mean number of retweets, used to approximate the spread and virality of policy-related messages.
- x_4 – **Average Followers:** the average number of followers of users posting relevant tweets, capturing the potential reach and influence of these discussions.
- x_5 – **Proportion of Trump Mentions:** the share of tweets mentioning "Trump," included due to his dominant role in fiscal policy debates during the 2020 election.
- x_6 – **Median State Income:** the average state-level median income, used to control for underlying regional economic conditions that may affect consumer confidence.
- x_7 – **Average Engagement Ratio:** calculated as the average of likes and retweets divided by the number of followers for each tweet. This normalizes engagement for user popularity and highlights content that resonates broadly regardless of account size.

- x_8 – **Average Sentiment:** sentiment scores of tweets (ranging from negative to positive), obtained via text analysis, which capture the emotional tone of online discussions about fiscal policy.

- x_9 – **Average State Funding:** the average government funding allocated per state in 2020 (in billions USD). This variable accounts for actual fiscal transfers that could affect both sentiment and economic conditions locally.

4 Visualization

The initial exploratory analysis reveals heterogeneity in both Twitter activity and consumer confidence across regions. Summary statistics indicate that regions with a higher volume of spending-related tweets tend to report higher consumer confidence scores. For instance, regions with intense Twitter discussions on fiscal policy displays an increase in consumer sentiment during the 2020 election cycle.

Figures 1 and 2 illustrate these findings graphically. Figure 1 presents a time-series plot comparing the daily volume of spending-related tweets with regional consumer confidence indices. Notably, peaks in social media activity often align with key policy events, such as the announcement of government relief packages. Figure 2 features a regional heatmap that visualizes the spatial distribution of digital sentiment. This map demonstrates that regions with more vibrant social media discussions exhibit higher average sentiment scores, which correlate with higher consumer confidence levels.

This choropleth map illustrates the sentiment intensity of Twitter engagement on the topic of government spending across U.S. states. The sentiment scores range from negative sentiment to positive sentiment, with colors transitioning from dark purple (low) to bright yellow (high). States such as Connecticut, Nebraska, and Massachusetts exhibit the highest engagement levels, indicating active and possibly polarized public discourse on fiscal policy. In contrast, states like South Dakota and West Virginia show minimal engagement, suggesting lower public interaction or interest on the subject within those regions. This visualization offers a geographic lens through which regional variations in online political discourse can be assessed.

Another map is made from a scraped dataset that displays the state funding in 2020. The darker shades represent higher funding levels. California received the largest share, followed by Texas, Florida, Pennsylvania, and New York. These states have large populations and significant infrastructure, which likely explains their higher allocations. In contrast, many central and Mountain West states, such as Wyoming, South Dakota, and Montana, received much less funding. These states tend to have smaller populations and fewer large-scale federal programs. The overall distribution reflects a strong correlation between population size and federal investment. However, some states with moderate populations, like Arizona and North Carolina, received less than expected, suggesting that factors beyond population, such as political influence, funding applications, or program eligibility may play a role. The map reveals clear regional differences and highlights the need to examine how funding decisions align with state-level needs and priorities. This can inform more equitable future allocation strategies.

Figure 1. Monthly Tweet Volume on Government Spending: Biden vs Trump (2020)

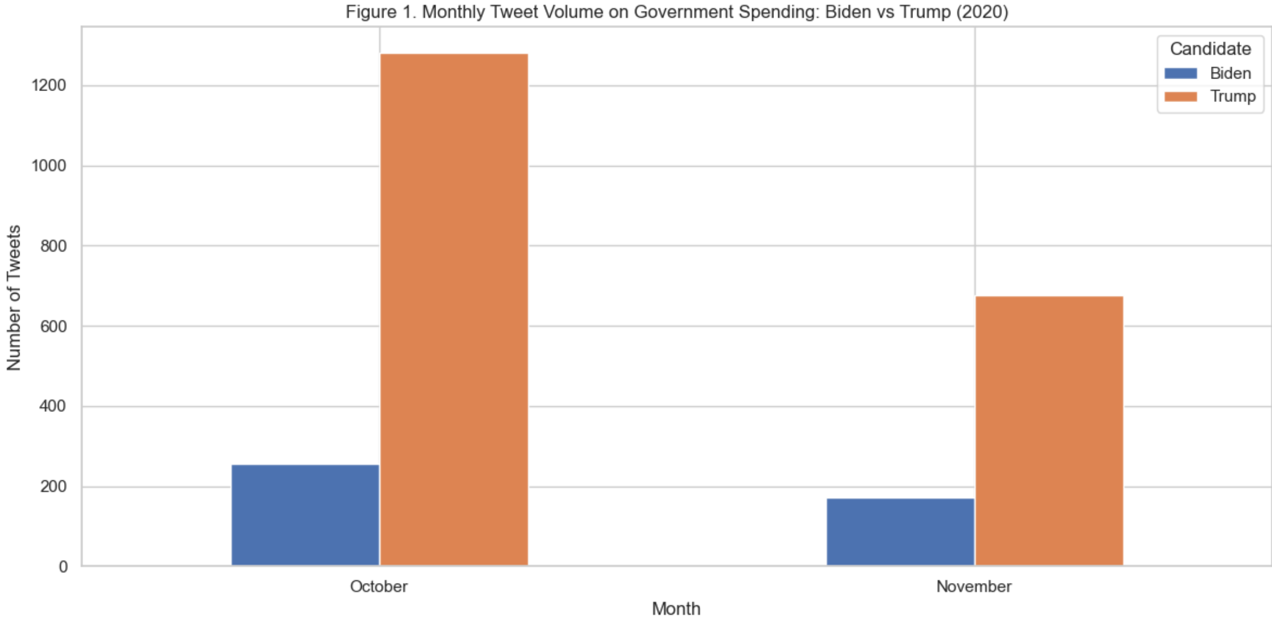
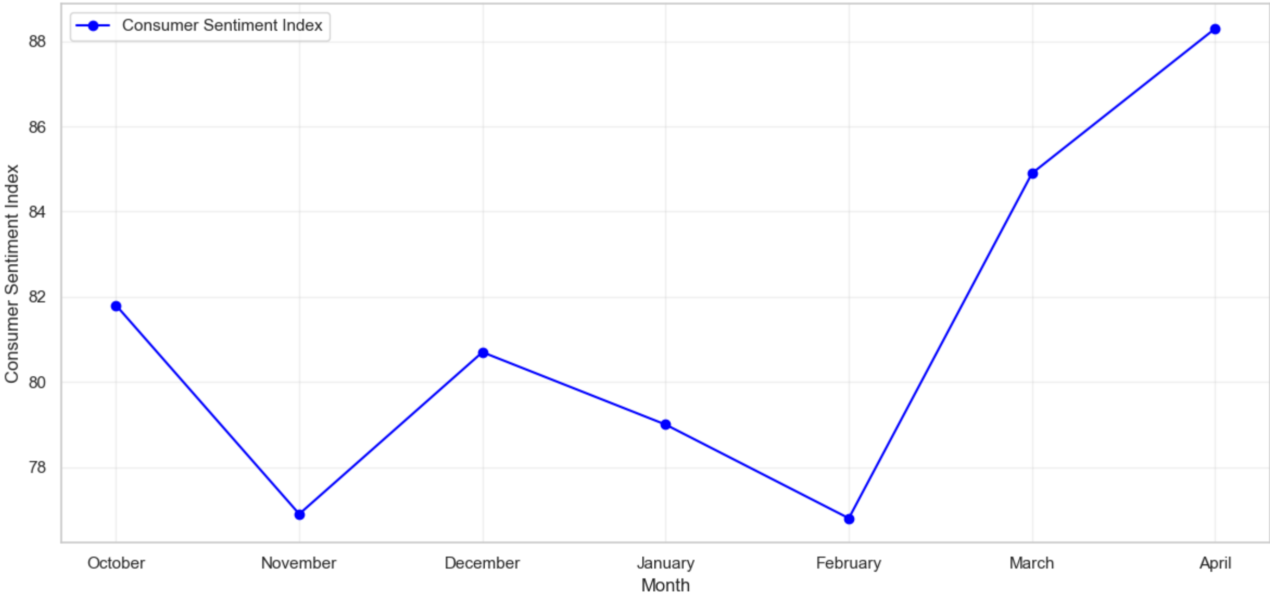
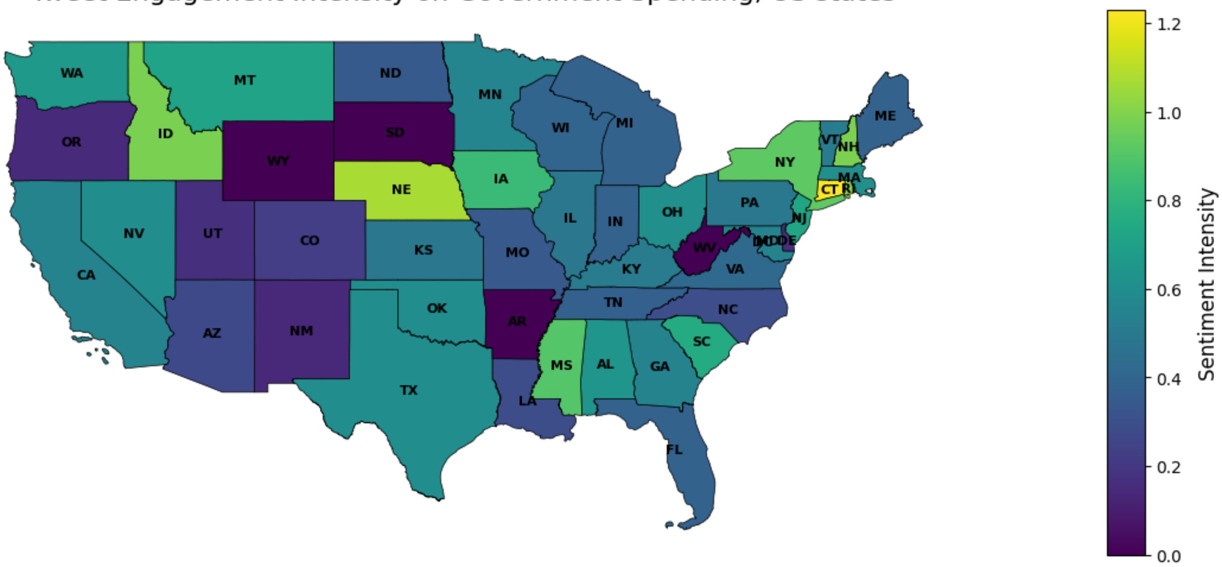


Figure 2. Monthly Consumer Sentiment Index Trends (Oct 2020 - Apr 2021)



Tweet Engagement Intensity on Government Spending, US states



Federal Funding by State (2020)

5 Regression Results

The goal is to address whether social media discussions about government spending affect consumer confidence. I focus on tweets posted during the 2020 US presidential election, examining their sentiment around fiscal policy to see how they might correlate with broader consumer confidence trends. While a linear relationship is possible where sentiment increases consumer confidence in a linear way, a non-linear relationship is relevant because very positive or very negative messages can have a bigger impact. During my analysis, shifts from low to moderate engagement appeared more influential than shifts from moderate to high which indicates diminishing returns. This made me include squared or interaction terms in some regressions to test for curvature in the data. I include several explanatory variables:

- Tweet Sentiment Score: A continuous measure of positivity/negativity for each tweet referencing government spending.

- Political Affiliation: A binary indicator (*istrump* or *isbiden*) for partisan effects. This is based on the assumption that messages from politicians often resonate differently with the public.

- Engagement Metrics: Number of likes, which reflects how widely a tweet is spread. Higher engagement could increase the tweet's influence.

- Geographic Controls: State or regional dummies, as local economic state may lead to different responses to spending discussions.

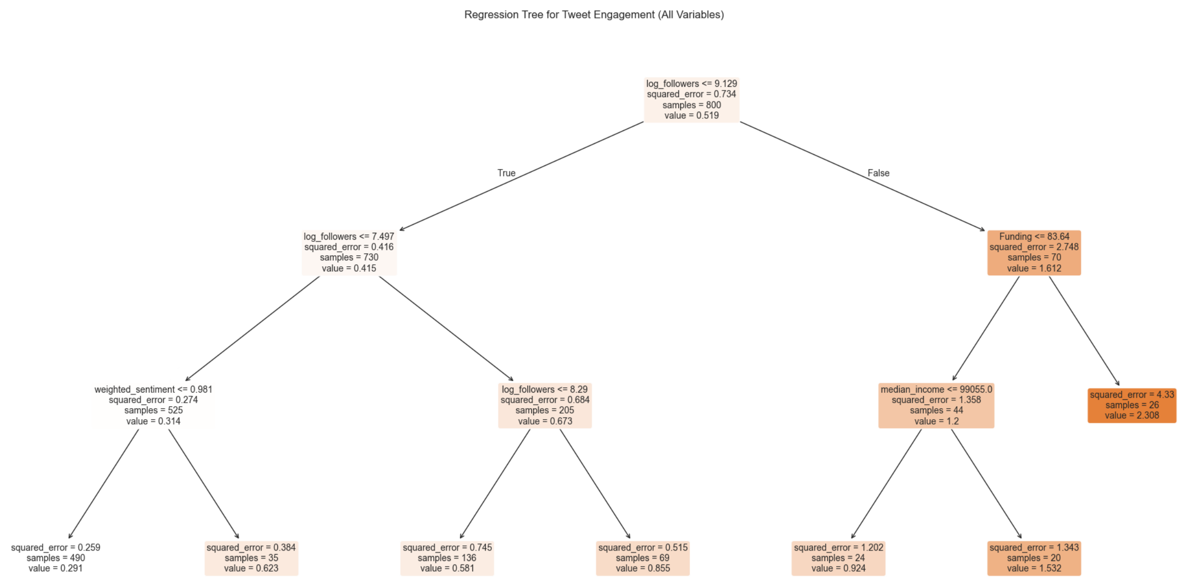
- Interaction Terms: For example, $trumplikesinteraction$ to test if the impact of likes on sentiment differs for Trump tweets.

Agents form expectations about fiscal policy not only from official statements but also from the media discussions. Including these variables helps to isolate how political preference and public engagement change sentiment around government spending.

The objective function for a regression tree with our variables can be written as:

$$\min_{\{R_j\}_{j=1}^{J}} \sum_{j=1}^{J} \sum_{i \in R_j} (y_i - \hat{y}R_j)^2$$

Where: R_j represents the j-th region (leaf node) y_i is the actual value of tweet engagement (log likes) for observation i $\hat{y}R_j$ is the predicted value for region R_j (mean of observations in that leaf) In plain terms, this objective function is trying to find the optimal way to split our data into regions (leaf nodes) that minimize the squared error between actual and predicted values. The regression tree recursively partitions our dataset using variables like sentiment scores, political affiliation (Trump/Biden), state funding levels, and user metrics to find homogeneous groups where tweet engagement is similar within each group but different between groups.



When all explanatory variables are used in a new regression tree, we see different splits and predictors compared to the simpler tree. Some social media factors like tweet sentiment and how often people discuss spending show up earlier in the new tree, which suggests they play a stronger role in explaining shifts in consumer confidence when all variables are considered. This implies that online discussions can be just as much influential as traditional economic measures when explaining changes in consumer confidence.

Economically the tree suggests that if social media sentiment about government spending is more positive, consumer confidence may rise and vice versa. This is consistent with earlier studies but shows how real-time online discussions can improve or worsen people's economic expectations.

As for the prediction errors, the new tree can capture nonlinear patterns like sudden jumps in confidence more easily than a standard regression, though there is still some unexplained variation. Despite a more flexible fit, we must remain cautious about overfitting, where the model becomes too tuned to the training data and loses accuracy with new data.

6 Conclusion

This study's findings show how social media discussions of fiscal policy can mirror and potentially influence public perceptions of the economy during a volatile political period such as the 2020 elections. By examining both the volume and sentiment of spending-related tweets alongside consumer confidence metrics, the analysis revealed that regions generating more robust online discussion tended to exhibit higher confidence levels. Moreover, emerging data on pandemic relief efforts: from stimulus checks to shifting economic policies, further highlights social media's role as a measure of public sentiment.

Nevertheless, the short time frame and the effects of a rapidly changing policy environment may go against drawing definitive causal conclusions. To improve these insights, future research should expand the dataset and leverage advanced analytical methods that can better disentangle online engagement from broader economic and political factors. In doing so, we can develop a more comprehensive understanding of the relationship between digital political conversations and real-world consumer attitudes.

References

[1] Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, *2*(1), 1–8.

[2] Choi, H., & Varian, H. (2012). Predicting the present with Google Trends. *Economic Record*, *88*, 2–9.

[3] Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance*, *66*(5), 1461–1499.

[4] Fisher, I. (1993). The debt-deflation theory of great depressions. *Econometrica*, *1*(4), 337–357.

[5] Garcia, D. (2013). Sentiment during recessions. *The Journal of Finance*, *68*(3), 1267–1300.

[6] Gao, P., & Jiang, W. (2017). Social networks, media, and investor behavior. *Management Science*, *63*(7), 2429–2451.

[7] Kim, D., Mikalef, P., & Pappas, I. O. (2014). Social media and consumer sentiment: Insights on the predictive power of social media. *International Journal of Information Management*, *34*(2), 260–270.

[8] Ludvigson, S. C. (2004). Consumer confidence and consumer spending. *Journal of Economic Perspectives*, *18*(2), 29–50.

[9] O'Connor, B., Balasubramanyan, R., Routledge, B. R., & Smith, N. A. (2010). From tweets to polls: Linking text sentiment to public opinion time series. *Proceedings of the International AAAI Conference on Web and Social Media*, *4*(1), 122–129.

[10] Oh, C., Park, Y., & Rao, H. R. (2015). Forecasting the diffusion of innovation considering competitive and cooperative relationships: A social network perspective. *Technological Forecasting and Social Change*, *95*, 108–120.

[11] Osterloh, S. (2018). How do politics affect economic sentiment? *Beiträge zur Jahrestagung des Vereins für Socialpolitik 2018: Digitale Wirtschaft — Session: Policy Uncertainty (No. F14-V2)*, ZBW - Leibniz-Informationszentrum Wirtschaft, Kiel, Hamburg.

[12] Öztürk, Z. (2024). Election and Wallet: The Effect of Election Periods and Consumer Expectations on Consumer Confidence with ARDL and QARDL Model Approach. *Bulletin of Economic Theory and Analysis*, *9*(3), 953–982.

[13] Preis, T., Moat, H. S., & Stanley, H. E. (2013). Quantifying trading behavior in financial markets using Google Trends. *Scientific Reports*, *3*(1), 1684.

[14] 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.

[15] Zhang, X., Fuehres, H., & Gloor, P. A. (2011). Predicting stock market indicators through Twitter "I hope it is not as bad as I fear". *Procedia - Social and Behavioral Sciences*, *26*, 55–62.