

# Economic Commentary

11.10.2025 | ISSN 2163-3738 | Number EC 2025-13



## Forecasting US Recessions in Real-Time Using Regional Economic Sentiment

Measures of regional economic sentiment, extracted from the Beige Book using natural language processing methods, consistently delivered reliable real-time forecasts of US recessions from the mid-1980s through the COVID-19 pandemic recession. Since then, recession risk probabilities have been choppy, with several false alarms. We attribute this unreliability to a post-2021 disconnect between measures of economic activity and the sentiment of business and community leaders.

Christian Garciga, James Mitchell

**Topics** [Business cycles](#), [Econometrics](#), [modeling](#), and [forecasting](#)

**DOI** [10.26509/frbc-ec-202513](https://doi.org/10.26509/frbc-ec-202513) [↗](#)

*The views authors express in Economic Commentary are theirs and not necessarily those of the Federal Reserve Bank of Cleveland or the Board of Governors of the Federal Reserve System. The series editor is Tasia Hane. This paper and its data are subject to revision; please visit [clevelandfed.org](https://clevelandfed.org) for updates.*

## Introduction

The Beige Book provides anecdotal and impressionistic summaries of the economic health of each of the Federal Reserve System's 12 Districts eight times a year.<sup>1</sup> These narratives reflect firsthand reports from business and community contacts collected by the Federal Reserve Banks along with information from other sources. In addition to a District-level report, each Beige Book provides a “national summary” summarizing the economic outlook for the US economy as a whole.

Since April 2024, Filippou et al. (2024) have been regularly producing quantitative estimates of regional economic sentiment extracted from the Beige Book. These estimates are publicly available at [clevelandfed.org/research/additional-research-resources](https://clevelandfed.org/research/additional-research-resources). The estimates that Filippou et al. (2024) produce use natural language processing (NLP) methods to quantify the economic sentiment expressed in these District-level and national summaries. In this *Economic Commentary*, we test how accurate these estimates have been at forecasting US recessions.

We find that regional economic sentiment yielded reliable real-time forecasts of US recessions from the mid-1980s until the end of 2021. The most accurate forecasts are obtained when an average is taken of economic sentiment across each of the 12 Federal Reserve Districts. But since 2021, recession risk probabilities have been more erratic. We attribute this to the historically unusual nature of the post-pandemic economic expansion and to changes in the drivers of regional economic sentiment.

## Constructing Regional Sentiment Indices

To construct our regional economic sentiment indices, we use BERT (bidirectional encoder representations from transformers), a deep learning NLP model developed by Google, to measure the sentiment of each sentence in the Beige Book. Specifically, we use a variant of BERT, called FinBERT (Huang et al., 2023), as this benefits from also being pretrained on financial text. The attraction of FinBERT is that it avoids the subjective use of judgment when defining sentiment.

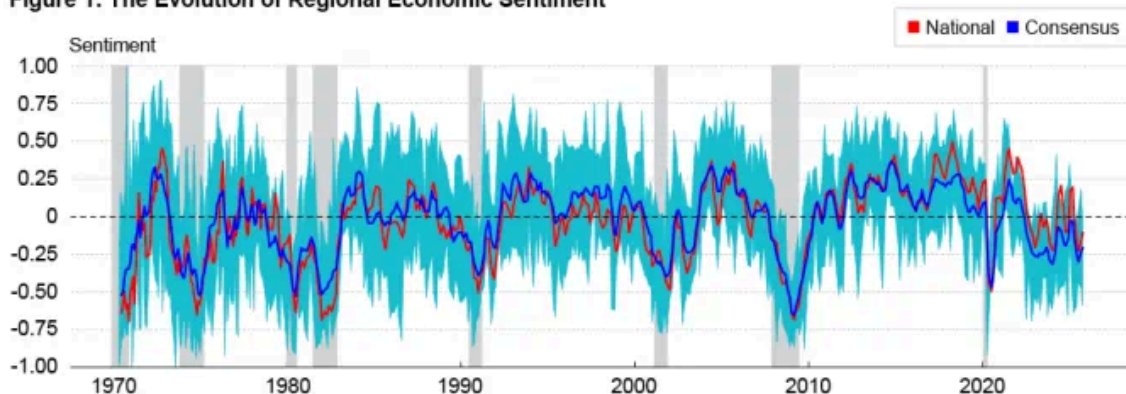
We start by using FinBERT to classify each sentence in the Beige Book as positive, neutral, or negative. Then we add up the number of positive sentences (either at the District-level or in the national economic summary) and subtract the number of negative sentences. Scaling this balance by the total number of positive and negative sentences, we arrive at an index of regional or national economic sentiment. This index takes a value between -1 (when sentiment is all negative), through 0 (when sentiment is neutral), to +1 (when sentiment is entirely positive). <sup>2</sup>

Figure 1 plots the latest set of FinBERT-based estimates of regional and national economic sentiment, as extracted from each successive Beige Book from 1970 through October 2025. The regional estimates are plotted as a shaded area capturing the lower and upper boundaries of the data ranges of the 12 measures of District-level sentiment extracted from each District report. Alongside this, as introduced in our previous *Economic Commentary* (Filippou et al., 2024), we plot the equal-weighted average of economic sentiment in each District, something we call “consensus” sentiment. Consensus sentiment can be contrasted with “national” sentiment, which is the measure of sentiment extracted by FinBERT from just the “national summary” in each Beige Book.

From Figure 1 we see that the sentiment indices generally rise and fall with the US business cycle, as characterized in gray bars by NBER-based turning points that differentiate expansionary from recessionary phases of the US business cycle. We also observe considerable variation in sentiment across the 12 Districts.

Anticipating a break in behavior that we discuss below, in recent years consensus sentiment has systematically been below national sentiment. Consensus sentiment has largely been in negative territory since mid-2022, averaging -0.2 from June 2022 through October 2025.

**Figure 1: The Evolution of Regional Economic Sentiment**



Sources: Federal Reserve Board Beige Book, NBER Business Cycle Dating Committee, and authors' calculations

Notes: "Consensus" sentiment is the equal-weighted average of sentiment in each of the 12 Reserve Bank Districts. "National" sentiment is from the "national summary" in the Beige Book, written by a designated Reserve Bank on a rotating basis. Shaded light-blue area captures the upper and lower boundaries of the data ranges capturing the 12 measures of District-level sentiment. Vertical shaded gray bars represent NBER recessions.

## Recession Forecast Accuracy

In Filippou et al. (2024), we showed that regional economic sentiment, as seen in Figure 1, contains statistically useful information about US business cycle phases, over and above the information contained in national sentiment. This finding—on the value-added of regional information even when focused on the US as a whole—is consistent with earlier research by Hamilton and Owyang (2011) and Owyang et al. (2015), which found that the business cycles of individual US states also frequently diverge from the national cycle. If these divergences are systematic, using regional data can enhance predictions of the US business-cycle phase. Additionally, even when Districts align with the US business cycle, individual Districts may experience recessions and expansions more intensely than the US as an aggregate. Including data from these Districts could offer a clearer indication of the current business-cycle phase, thereby improving nowcasts and forecasts of the national business cycle.

We extend our previous *Economic Commentary* to test whether the historical, or “in-sample,” correlations between our regional economic indices and US business-cycle phases, as seen in Figure 1, hold up in real time, or “out-of-sample.” This is an important test. It is well-known from the forecasting literature that strong in-sample fit between a set of predictors and the target variable does not

necessarily translate into reliable out-of-sample forecasts.<sup>3</sup> So-called temporal instabilities—that is, changes in relationships between variables—often plague macroeconomic forecasting (Rossi, 2021).

The specific forecasting problem that we consider is to predict, using our estimates of sentiment from the Beige Book released in month  $t$ , whether the US economy will be in a recession at time  $(t+h)$ . Given that the NBER's Business Cycle Dating Committee tends to publish its recession dates at a lag, often many months after the beginning of a recession, it is of interest to backcast ( $h<0$ ) expansions/recessions as well as to nowcast ( $h=0$ ) and forecast ( $h>0$ ) future ones. We focus on recession forecasts  $h=-1, h=0, h=1$  months ahead.

To mimic real-time use of our text-based estimates of regional economic sentiment, we undertake the following out-of-sample simulation. We start at December 31, 1984, and we use data known at this point in time, specifically Beige Book data from 1970 through the end of 1984, to estimate binary (logit) regressions relating NBER-dated US recessions  $h=-1, h=0$ , and  $h=1$  months ahead to District- and national-level sentiment. We use these three estimated models and the sentiment estimates from the January 1985 Beige Book to produce recession probability backcasts for December 1984 ( $h=-1$ ), nowcasts for January 1985 ( $h=0$ ), and forecasts for February 1985 ( $h=1$ ). We then update our estimation sample with the January 1985 Beige Book and repeat the exercise, producing forecasts upon the release of the next Beige Book in March 1985. This process of recursively updating the sample and producing new recession probability forecasts carries on until we use information up to and including the September 2025 Beige Book to backcast the probability of a recession in August 2025 ( $h=-1$ ), nowcast ( $h=0$ ) the probability of one in September 2025, and forecast the probability of a recession in October 2025 ( $h=1$ ).

To tease out the best way of using our regional economic sentiment indices when forecasting US recessions in real time, we delineate five forecasting models that use different sets of the regional and national sentiment indices. Specifically, we compare the performance of five models with the following variables as predictors:

Model 1: National sentiment and the 12 District-level sentiment indices

Model 2: 12 District-level sentiment indices

Model 3: Consensus sentiment alone

Model 4: National sentiment alone

Model 5: Consensus and national sentiment

Comparing forecast accuracy across these five models lets us isolate whether there is value added in District-level sentiment and, if so, what the best way to exploit it is. Is it best to let the data determine the weights on each District, as in models 1 and 2, or to use the equal-weighted average as characterized by consensus sentiment? A large body of literature finds that equal-weighted averages often outperform weighted combinations in out-of-sample forecasting exercises (Smith and Wallis, 2009). Comparing across models 3, 4, and 5 lets us test whether there is information in consensus sentiment, over and above that contained in national sentiment, that is helpful when forecasting US recessions.

To evaluate the ability of the five models to separate expansionary and recessionary phases of the business cycle, we follow a recent but growing literature in economics (Berge and Jordà, 2011) and use the area under the receiver operating characteristic curve (AUC). The receiver operating characteristic curve describes the relationship between the “false positive” rate and the “true positive” rate. Specifically, it lets us see how frequently the model incorrectly predicts a recession (false positive) and how frequently it calls one correctly (true positive). These rates are dependent on the threshold used to convert the recession probability forecast (a value between 0 percent and 100 percent) into a binary forecast (recession or expansion). In general, there is a trade-off when selecting this threshold. Lowering the threshold tends to increase both the true positive and false positive rates. This means more actual recessions are correctly identified, but at the cost of more false alarms. Raising the threshold tends to decrease both true positive and false positive rates. This results in fewer false alarms, but potentially at the cost of missing some actual recessions. The AUC provides a single summary measure of predictive accuracy across all possible thresholds. An AUC of unity indicates a perfect classifier. An AUC of 0.5 suggests that the model is no better than a random guess, equivalent to the outcome of a coin toss. The better our five models can forecast recessions, the closer their AUC values are to one.

It is also helpful to isolate forecast performance at specific points on the receiver operating characteristic curve to examine the associated true positive and false positive rates. We do so by reporting Youden’s index. Youden’s index finds the optimal threshold on the receiver operating characteristic curve by maximizing the difference between the true positive rate and the false positive rate.<sup>4</sup> Again, the closer the Youden index to unity, the better the forecast.



**Table 1: Evaluating the Out-of-Sample Accuracy of the Recession Probability Forecasts from the Five Models: AUC Values and Youden's Index**

Model	AUC			Youden's Index		
	h=-1	h=0	h=1	h=-1	h=0	h=1
M1: National + 12 Districts	0.99	0.95	0.93	0.96	0.82	0.79
M2: 12 Districts	0.99	0.95	0.93	0.96	0.84	0.79
M3: Consensus	0.99	0.96	0.95	0.97	0.84	0.82
M4: National	0.98	0.94	0.91	0.87	0.79	0.74
M5: National + Consensus	0.99	0.96	0.95	0.97	0.84	0.82

Sources: Federal Reserve Board Beige Book, NBER Business Cycle Dating Committee, and authors' calculations

Notes: Out-of-sample period is January 1985 through September 2025. Model 1 (M1) through model 5 (M5) are as explained in text.

Table 1 reports the AUC values and Youden's indices from the five models. We draw out four main takeaways. First, as we should expect, forecast accuracy, as measured by both metrics, improves the further into the past we look. AUC and Youden's index values are closer to unity for the backcasts than for the forecasts. Second, forecasting US recessions using only information from the national economic summary (model 4) consistently produces worse forecasts than the other four models. This demonstrates that there is value added in District-level sentiment indices. Third, the best way to harness this regional information is to use consensus sentiment. While the gains are modest, consistently we see the "forecast combination puzzle" (Smith and Wallis, 2009), that is, equal-weighted combinations of District-level sentiment (as seen in model 3) yield higher AUC and Youden's index values than using the data to estimate the weights on each District (model 2). Fourth, since models 3 and 5 are equally accurate, we conclude that consensus sentiment encompasses (Chong and Hendry, 1986) national sentiment. There is no value added in national sentiment over and above the information already in consensus sentiment. Model 3 is preferred to model 5 because it is the simpler model.

Overall, the AUC values in Table 1 suggest that regional economic sentiment delivers extremely reliable recession forecasts. An AUC value of 0.96 at  $h=0$  using consensus sentiment (M3) indicates that Consensus sentiment has excellent classification ability. This performance compares well with the accuracy of alternative indices of current business conditions, as evaluated in Berge and Jordà (2011), such as the Chicago Fed national activity index and the Aruoba, Diebold, and Scotti (ADS) business conditions index maintained by the Federal Reserve Bank of Philadelphia.<sup>5</sup> <sup>6</sup> It also compares well with recession-forecasting models that use financial indicators; see Burke and Nelson (2025).

## Real-Time Recession Forecasts

Having established that, on average over time, the most accurate recession forecasts come from using consensus economic sentiment, Figure 2 plots the recession probability nowcasts ( $h=0$ ) from model 3. Consistent with Table 1, the overall impression from Figure 2 is again that consensus

sentiment provides reliable estimates of US recessions. The forecasted probabilities of recession clearly rise in periods subsequently classified as a recession by the NBER.

As discussed, when using probability forecasts for classification, there is always a question of what threshold to use. The optimal threshold value above which it is “optimal” to call a recession when nowcasting is 26 percent.<sup>7</sup> Using this, we see one false positive in the run-up to the 2001 recession. We also see a second short series of false alarms, when the recession probabilities remained elevated as the economy emerged from the global financial crisis in 2009. While these were both “false” alarms, they were right around the start and end dates of recessions.

However, since the end of 2021, the recession risk probabilities in Figure 2 have been higher and choppy. But the US economy has not entered a recession since 2020, based on currently available information from the NBER’s Business Cycle Dating Committee.<sup>8</sup>

We see three main false alarms in Figure 2, when the recession probabilities markedly rose above the optimal threshold. These three spikes occurred in November 2022, October/November 2023, and April 2025.<sup>9</sup>

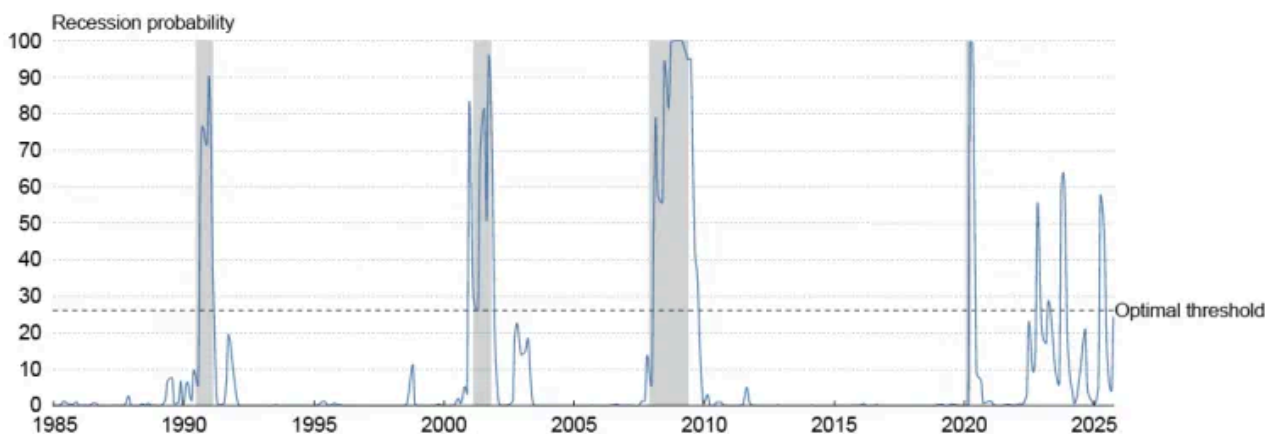
The spike in April 2025 has an obvious cause, confirmed when reading this Beige Book (published on April 23): tariff and international trade policy uncertainties. The package of new tariffs announced on April 2, 2025, featured heavily in the April 2025 Beige Book. This is reflected by economic sentiment dropping sharply in each of the 12 Districts.<sup>10</sup>

But the earlier two spikes in Figure 2 do not appear to have such specific causes. Rather, we interpret them as symptomatic of a growing post-pandemic disconnect between economic sentiment and business-cycle variables of the sort used by the NBER when dating recessions. This disconnect is evidenced when, in an illustrative exercise, we relate consensus economic sentiment to five macroeconomic variables. Specifically, we regress consensus economic sentiment on the prevailing US GDP growth rate, unemployment rate, inflation rate, federal funds rate, and an indicator of economic uncertainty.<sup>11</sup> We include GDP growth in the regression as a proxy for the various measures of aggregate real economic activity that the NBER consults when classifying recessions.<sup>12</sup> The unemployment rate and the inflation rate comprise the constituents of the so-called misery index proposed by Arthur Okun in the 1970s as a measure of economic distress. The federal funds rate captures the stance of monetary policy and is a proxy for wider borrowing costs. Finally, we consider the text-based economic policy uncertainty index of Baker et al. (2016) to capture any link between sentiment in the Beige Book and wider public discussions of economic policy uncertainty, as reflected in newspaper coverage.

Table 2 shows that, prior to 2021, GDP growth and inflation were statistically significant drivers of Beige Book sentiment, with stronger GDP growth associated with improving sentiment and, by contrast, higher inflation depressing sentiment.<sup>13</sup> But since the pandemic, neither GDP growth nor inflation explain movements in sentiment in a statistically significant manner. This helps us understand why sentiment has become a less reliable forecaster of recessions, given the importance of GDP—and related indicators of aggregate real economic activity—to the NBER Business Cycle Dating Committee. Echoing the results in Bolhuis et al. (2024) for consumer sentiment, Table 2 provides some evidence that sentiment has become more sensitive to interest rates since the pandemic. The estimated coefficient on the federal funds rate in Table 2 quintuples in size after 2021. In other words, post-pandemic-recession movements in GDP growth and the misery index (alone) cannot explain why consensus economic sentiment has been so low. Like consumers, business and community leaders have felt the “vibecession.”<sup>14</sup>

We can also begin to understand the post-pandemic-recession volatility of the recession probability forecasts in Figure 2 by observing from Table 2 that Beige Book sentiment is strongly negatively correlated with economic policy uncertainty. Measures of uncertainty, including the economic policy uncertainty index that we focus on, have been especially elevated and volatile since 2021.<sup>15</sup> Heightened uncertainty may capture various forces, including public discussions of whether raising the federal funds rate would result in a soft landing, the effects of changes in trade and other public policies, and the generally unusual nature of this expansion.<sup>16</sup> By averaging sentiment across various topics discussed in the Beige Book, we see that recently our regional sentiment indices may be providing noisier estimates of the business cycle compared to during periods when sentiment was more closely aligned with traditional economic variables such as GDP that are typically used to classify recessions.

**Figure 2: Real-Time Nowcast Probabilities of a US Recession Using Consensus Sentiment**



Sources: Federal Reserve Board Beige Book, NBER Business Cycle Dating Committee, and authors' calculations

Note: Vertical shaded bars represent NBER recessions.



Table 2. The Drivers of Consensus Economic Sentiment

	1970:M1–2020:M12	2021:M1–2025:M7
GDP growth	0.021** (0.009)	-0.009 (0.013)
Unemployment rate	-0.011 (0.013)	-0.014 (0.080)
Inflation	-0.022** (0.009)	-0.048 (0.035)
Federal funds rate	-0.016*** (0.005)	-0.084*** (0.027)
Economic policy uncertainty	-0.091** (0.036)	-0.049*** (0.009)
R <sup>2</sup>	0.45	0.67

Sources: Federal Reserve Board Beige Book, NBER Business Cycle Dating Committee, and authors' calculations




Note: HAC standard errors with six lags (1970:M1–2020:M12) and four lags (2021:M1–2025:M7) in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .












## Conclusion

This *Economic Commentary* shows how NLP methods can be used to produce recession-risk forecasts from the Beige Book's narrative on economic conditions in the 12 Federal Reserve Districts and the nation as a whole. Our out-of-sample evidence finds that the most accurate real-time forecasts of US recessions are obtained, not from national economic sentiment, but by taking an equal-weighted average of District-level economic sentiment. For more than 30 years, from the mid-1980s through the COVID-19 pandemic recession, these forecasts proved to be extremely reliable.

But since the pandemic recession, regional economic sentiment has whipsawed and, abusing Paul Samuelson's famous quip about the stock market, has predicted three of the last zero recessions. We attribute this newfound unreliability to the historically unusual nature of the post-pandemic-recession economic expansion. Our nowcast model predicts that there is a 24 percent chance that the US economy was in a recession in October 2025. But recent swings in our recession probability forecasts reinforce the importance of linking the narrative or story behind the forecasts, as provided by the Beige Book text, to the forecasts themselves. The weaker this link, the less confidence we should have in the quantitative forecasts themselves.

## References

- Baker, Scott R., Nicholas Bloom, and Steven J. Davis. 2016. "Measuring Economic Policy Uncertainty." *Quarterly Journal of Economics* 131(4): 1593–1636. [doi.org/10.1093/qje/qjw024](https://doi.org/10.1093/qje/qjw024) .
- Berge, Travis J., and Òscar Jordà. 2011. "Evaluating the Classification of Economic Activity into Recessions and Expansions." *American Economic Journal: Macroeconomics* 3(2): 246–277. [doi.org/10.1257/mac.3.2.246](https://doi.org/10.1257/mac.3.2.246) .
- Bolhuis, Marijn A., Judd N.L. Cramer, Karl Oskar Schulz, and Lawrence H. Summers. 2024. "The Cost of Money is Part of the Cost of Living: New Evidence on the Consumer Sentiment Anomaly." NBER Working Paper 32163. National Bureau of Economic Research. [doi.org/10.3386/w32163](https://doi.org/10.3386/w32163) .

- Burke, Mary A., and Nathaniel R. Nelson. 2025. "The Beige Book's Value for Forecasting Recessions." *Current Policy Perspectives* No. 25-15. Federal Reserve Bank of Boston. [bostonfed.org/publications/current-policy-perspectives/2025/beige-book-for-forecasting-recessions.aspx](https://bostonfed.org/publications/current-policy-perspectives/2025/beige-book-for-forecasting-recessions.aspx) .
- Caldara, Dario, Matteo Iacoviello, Patrick Molligo, Andrea Prestipino, and Andrea Raffo. 2020. "The Economic Effects of Trade Policy Uncertainty." *Journal of Monetary Economics* 109: 38–59. [doi.org/10.1016/j.jmoneco.2019.11.002](https://doi.org/10.1016/j.jmoneco.2019.11.002) .
- Chong, Yock Y., and David F. Hendry. 1986. "Econometric Evaluation of Linear Macro-economic Models." *Review of Economic Studies* 53(4): 671–690. [doi.org/10.2307/2297611](https://doi.org/10.2307/2297611) .
- Clark, Todd E. 2004. "Can Out-of-Sample Forecast Comparisons Help Prevent Overfitting?" *Journal of Forecasting* 23(2): 115–139. [doi.org/10.1002/for.904](https://doi.org/10.1002/for.904) .
- Filippou, Ilias, Christian Garciga, James Mitchell, and My T. Nguyen. 2024. "Regional Economic Sentiment: Constructing Quantitative Estimates from the Beige Book and Testing Their Ability to Forecast Recessions." *Economic Commentary*, no. 2024-08. [doi.org/10.26509/frbc-ec-202408](https://doi.org/10.26509/frbc-ec-202408).
- Gascon, Charles S., and Joseph Martorana. 2025. "Quantifying the Beige Book's 'Soft' Data." *On the Economy*. Federal Reserve Bank of St. Louis. [stlouisfed.org/on-the-economy/2025/jan/quantifying-beige-books-soft-data](https://stlouisfed.org/on-the-economy/2025/jan/quantifying-beige-books-soft-data) .
- Granger, Clive W.J., and M. Hashem Pesaran. 2000. "Economic and Statistical Measures of Forecast Accuracy." *Journal of Forecasting* 19(7): 537–560. [doi.org/10.1002/1099-131X\(200012\)19:7<537::AID-FOR769>3.0.CO;2-G](https://doi.org/10.1002/1099-131X(200012)19:7<537::AID-FOR769>3.0.CO;2-G) .
- Hamilton, James D., and Michael T. Owyang. 2012. "The Propagation of Regional Recessions." *The Review of Economics and Statistics* 94(4): 935–947. [doi.org/10.1162/REST\\_a\\_00197](https://doi.org/10.1162/REST_a_00197) .
- Huang, Allen H., Hui Wang, and Yi Yang. 2023. "FinBERT: A Large Language Model for Extracting Information from Financial Text." *Contemporary Accounting Research* 40(2): 806–841. [doi.org/10.1111/1911-3846.12832](https://doi.org/10.1111/1911-3846.12832) .
- Koop, Gary, Stuart McIntyre, James Mitchell, and Aubrey Poon. 2022. "Reconciled Estimates of Monthly GDP in the United States." *Journal of Business and Economic Statistics* 41(2): 563–577. [doi.org/10.1080/07350015.2022.2044336](https://doi.org/10.1080/07350015.2022.2044336) .
- Owyang, Michael T., Jeremy Piger, and Howard J. Wall. 2015. "Forecasting National Recessions Using State-Level Data." *Journal of Money, Credit and Banking* 47(5): 847–866. [doi.org/10.1111/jmcb.12228](https://doi.org/10.1111/jmcb.12228) .
- Rossi, Barbara. 2021. "Forecasting in the Presence of Instabilities: How We Know Whether Models Predict Well and How to Improve Them." *Journal of Economic Literature* 59(4): 1135–1190. [doi.org/10.1257/jel.20201479](https://doi.org/10.1257/jel.20201479) .
- Smith, Jeremy, and Kenneth F. Wallis. 2009. "A Simple Explanation of the Forecast Combination Puzzle." *Oxford Bulletin of Economics and Statistics* 71(3): 331–355. [doi.org/10.1111/j.1468-0084.2008.00541.x](https://doi.org/10.1111/j.1468-0084.2008.00541.x) .

## Endnotes

1. Through much of the 1970s the Beige Book (then called the Redbook) was published more than eight times per year. We exploit these additional reports in our analysis below. [Return to 1](#)
2. Specifically, our measure of economic sentiment is computed as:  $\frac{n_t^{Positive} - n_t^{Negative}}{n_t^{Positive} + n_t^{Negative}}$ , where  $n_t^{Positive}$  represents the total number of positive sentences published in month  $t$ , and  $n_t^{Negative}$  represents the total number of negative sentences published in month  $t$ . Throughout this *Economic Commentary*, when we refer to the Beige Book in month  $t$ , we are referencing its publication date. Occasionally, the Beige Book is published at the

beginning of the month following the month indicated in its title.

[Return to 2](#)

3. For example, see Clark (2004). [Return to 3](#)
4. Granger and Pesaran (2000) discuss the links between the Kuipers score (the difference between the true positive and false positive rates), widely used in meteorology, and market timing tests, seen in finance. They explain how the metrics can be used to assess the economic value of forecasting models. [Return to 4](#)
5. These indices, and details of their construction, are available at [chicagofed.org/research/data/cfnai/current-data](https://chicagofed.org/research/data/cfnai/current-data) [↗](#) and [philadelphiafed.org/research-and-data/real-time-center/business-conditions-index](https://philadelphiafed.org/research-and-data/real-time-center/business-conditions-index) [↗](#). [Return to 5](#)
6. It is also noteworthy that if we smooth the regional sentiment indices, by taking a three-month moving average prior to estimating the logit models, forecast accuracy is systematically worse than when we use the raw sentiment indices evaluated in Table 1. [Return to 6](#)
7. This value is benefitting from look-ahead bias, as Youden's index is estimated (ex post) over the out-of-sample window. [Return to 7](#)
8. These post-pandemic-recession forecasting failures are reflected in the evaluation metrics reported in Table 1. The AUC value (Youden index) rises from 0.96 (0.84) in Table 1 to 0.99 (0.91) when we reevaluate the nowcasts from M3 on an out-of-sample window (1985:M1 through 2019:M12) that ends before the COVID-19 pandemic. [Return to 8](#)
9. It is worth noting that each Beige Book is based on information collected in around a six-week window that ends about a week before Beige Book publication. [Return to 9](#)
10. External measures of trade policy uncertainty, such as those that count the frequency of trade policy and uncertainty terms across major newspapers (Caldara et al., 2020), also shot up to unprecedentedly high values in April 2025; see [policyuncertainty.com/trade\\_cimpr.html](https://policyuncertainty.com/trade_cimpr.html) [↗](#). [Return to 10](#)
11. We consider the monthly “true” real GDP estimates (rolling quarter-on-quarter annualized percentage changes) from Koop et al. (2024), the 12-month PCE inflation rate, unemployment rate, effective federal funds rate, and economic policy uncertainty index of Baker et al. (2016) divided by 100. We focus on relating sentiment in month  $t$  to time  $t$  values of the macro indicators and ignore any timing issues caused by publication lags associated with official data. Regressions include an intercept. [Return to 11](#)
12. See [nber.org/research/business-cycle-dating](https://nber.org/research/business-cycle-dating) [↗](#). [Return to 12](#)
13. The regressions in Table 2 describe correlations, not causation, per se. [Return to 13](#)
14. See [brookings.edu/articles/the-paradox-between-the-macroeconomy-and-household-sentiment/](https://brookings.edu/articles/the-paradox-between-the-macroeconomy-and-household-sentiment/) [↗](#). [Return to 14](#)
15. More targeted uncertainty measures, such as the monetary policy uncertainty index of Baker, Bloom, and Davis (2016), have been twice as high after the pandemic recession as their average value between 1985 and 2020. [Return to 15](#)
16. Gascon and Martorana (2025) highlight the importance of physical (natural disasters) and political shocks as important noncyclical drivers of Beige Book sentiment. [Return to 16](#)

## Suggested Citation

Garciga, Christian, and James Mitchell. 2025. “Forecasting US Recessions in Real-Time Using Regional Economic Sentiment.” Federal Reserve Bank of Cleveland, Economic Commentary 2025-13. <https://doi.org/10.26509/frbc-ec-202513> [↗](#)

This work by [Federal Reserve Bank of Cleveland](#) is licensed under Creative Commons Attribution-NonCommercial 4.0 International 