ECONOMIC COMMENTARY

Regional Economic Sentiment: Constructing Quantitative Estimates from the Beige Book and Testing Their Ability to Forecast Recessions

Ilias Filippou, Christian Garciga, James Mitchell, and My T. Nguyen

We use natural language processing methods to quantify the sentiment expressed in the Federal Reserve's anecdotal summaries of current economic conditions in the national and 12 Federal Reserve District-level economies as published eight times per year in the Beige Book since 1970. We document that both national and District-level economic sentiment tend to rise and fall with the US business cycle. But economic sentiment is extremely heterogeneous across Districts, and we find that national economic sentiment is not always the simple aggregation of District-level sentiment. We show that the heterogeneity in District-level economic sentiment can be used, over and above the information contained in national economic sentiment, to better forecast US recessions.

Sentiment and stories about the economy matter. Economic research and the popular press often emphasize the importance of sentiment, of John Maynard Keynes' "animal spirits," in driving macroeconomic outcomes. In turn, Robert Shiller's 2019 book *Narrative Economics* describes economic fluctuations as driven by swings in popular economic narratives. In setting monetary policy, the Federal Open Market Committee (FOMC) has also long been aware of the importance of gathering local and national economic intelligence, including on sentiment about the economy. The FOMC consults a wealth of information about the US economy as a whole and for its regions. This information includes the Beige Book, a unique and timely source of qualitative evidence on the state of the Federal Reserve System's 12 District economies in the United States.

The Beige Book has been published eight times per year since 1970, with a new issue released about two weeks before each FOMC meeting. It contains reports from each of the 12 Federal Reserve Banks that draw on anecdotal and impressionistic information on economic conditions in each District gathered through firsthand reports from Bank and branch directors and interviews with key business contacts and other sources. On top of these 12 regional reports, the Beige Book provides a national summary that draws on the 12 District reports.[1] The national summary is prepared by a designated Reserve Bank on a rotating basis.

Ilias Filippou is a visiting assistant professor of finance at Washington University in St. Louis. Christian Garciga is an economic analyst at the Federal Reserve Bank of Cleveland. James Mitchell is a vice president at the Federal Reserve Bank of Cleveland. My T. Nguyen is a postdoctoral research associate in finance at Washington University in St. Louis.

In this *Economic Commentary*, we use natural language processing methods to quantify and then measure the information content of the sentiment expressed in the Beige Book, both about the nation as a whole and for the 12 District economies. The methods we use automate the quantification of economic sentiment from the Beige Book text. This design contrasts with more subjective, judgment-informed efforts to process the Beige Book narrative, as used in previous studies such as Balke and Petersen (2002) and Zavodny and Ginther (2005). We show that while sentiment across the 12 Districts comoves, with optimism rising during national expansions and falling during national recessions, there is considerable heterogeneity across the Districts. District-level sentiment is not always equally dispersed around national-level sentiment. Since the 2020 recession, our measure of national sentiment has been more optimistic about the performance of the national economy than the aggregation of sentiment in each of the 12 Districts. Historically, exploiting the heterogeneity in our new District-level sentiment indices delivers better forecasts of US recessions than conditioning on national sentiment alone.

Text as Data: Sentiment Analysis Using BERT

Given their increasing availability in digitized form, text-based data are now a popular resource when modeling and forecasting in macroeconomics. Sentiment and confidence have long been measured using qualitative data that draw on surveys asking people whether they think the economy is improving, staying the same, or getting worse. Such sentiment measures have been shown to have predictive power for macroeconomic outcomes, even when controlling for other factors.[2] A specific type of sentiment—uncertainty—is often measured by counting the number of times "uncertainty" or related words appear in a body of text and then measuring the number's evolution over time. A leading example of this is the economic policy uncertainty index of Baker, Bloom, and Davis (2016), rises in which have been found to have negative macroeconomic effects.

Rather than rely on this type of "bag of words" approach to quantify sentiment in the Beige Book, we follow recent research (such as Gorodnichenko, Pham, and Talavera, 2023) and turn to BERT (bidirectional encoder representations from transformers).[3] Developed by researchers at Google in 2018, BERT is a deep learning model used for natural language processing. It focuses on sequences of words rather than simply counting particular words in isolation, as in the economic policy uncertainty index. Pretrained on a corpus of more than 3.3 million words, with over 110 million parameters (Devlin et al., 2019), BERT avoids the subjective use of judgment in choosing the dictionary used to define, in our context, sentiment. Importantly, by using surrounding text, rather than simply reading from left to right, BERT aims to establish the context of the text and thus infer the meaning of language that might otherwise be ambiguous.

An attraction of BERT is that, because it is pretrained, it can be used out of the box. To benefit further from also being trained to analyze the sentiment of financial text, we use a variant of BERT: FinBERT (Huang, Wang, and Yang, 2023).[4] We apply FinBERT to the Beige Book at the sentence level. Specifically, for a given Beige Book, we partition its text into sentences and apply FinBERT, which then classifies sentences as positive, neutral, or negative. We measure the tone of each Beige Book as the difference between the number of positive sentences and negative sentence divided by the combined number of positive and negative sentences in that Beige Book. Specifically, the measure takes the following form:

$$\text{Tone}_t^{\text{FinBERT}} = \frac{n_t^{\text{Positive}} - n_t^{\text{Negative}}}{n_t^{\text{Positive}} + n_t^{\text{Negative}}} \tag{1}$$

where n_t^{Negative} represents the total number of negative sentences in the Beige Book corpus published in month t, and n_t^{Positive} represents the total number of positive sentences in the corpus published in month t. The sentiment measure in equation (1) is bounded by -1 and 1, with higher values indicating more positive sentiment. We use tone as a synonym for sentiment, as in papers such as Thorsrud (2020).

Following Barsky and Sims (2012) and Angeletos, Collard, and Dellas (2018), our measure of sentiment in equation (1) can capture both fundamental information (news) about the current and future economy or reflect a causal channel from sentiment (Keynes' animal spirits) to macroeconomic outcomes as optimism and pessimism rise and fall. We neither attempt to disentangle these two contrasting interpretations of sentiment nor establish causality.[5] Instead, our focus is establishing (reduced-form) statistical facts about District-level and national sentiment in the Beige Book and examining their predictive power for US recessions. In so doing, our sentiment measure is an aggregated measure in the sense that it quantifies sentiment across different areas of the economy, such as overall economic activity, labor markets, prices, and different sectors.

Our textual analysis focuses on using FinBERT to measure national and District-level sentiment in the 468 successive Beige Books starting from May 1970 through March 2024 that comprise the entire historical database. We access the Beige Book back to 1970 from the public archive maintained by the Federal Reserve Bank of Minneapolis. Over this sample period, the number of words in the Beige Book has increased. The average number of words in the Beige Book is around 1,300. The Beige Book provides a description of economic activity in each of the 12 Districts, as prepared by the respective regional Reserve Bank, and a national summary that is prepared by one of the regional Reserve Banks on a rotating basis. Each of these District-level and national reports starts with a summary of economic conditions before turning to a more specific discussion of economic conditions, including in the labor market, wages and prices, and specific industries of importance in that District.

Features of National and District-Level Sentiment

Figure 1 plots our FinBERT-based measure of national sentiment as extracted from the national summary in each successive Beige Book from 1970 through the latest Beige Book. Alongside this, we plot a shaded area capturing the upper and lower boundaries of the data ranges capturing the 12 measures of District-level sentiment extracted from each District report in the Beige Book. Figure 1 also plots the equal-weighted average of sentiment in each District, an average which we call "consensus" sentiment. In macroeconomic forecasting, it is common to refer to equal-weighted averages of different forecasts as consensus forecasts, and so we adopt this nomenclature here.

We emphasize two key features of Figure 1.[6] First, focusing on the national sentiment index (in red) we see a clear relationship between rises and falls in sentiment and business cycle turning points for the United States as a whole, as identified by the National Bureau of Economic Research (NBER) and marked as vertical gray bars in Figure 1. As expected, national sentiment falls in recessions.[7] But Figure 1 reveals that sentiment did not drop as deeply in the Gulf War recession of July 1990 to March 1991 or the dotcom recession of March 2001 to November 2001. This finding fits with the quantitative fact that both of these recessions were characterized by smaller peak-to-trough declines in GDP and smaller rises in the unemployment rate than in the other recessions since 1973. The December 1969 to November 1970 recession at the beginning of our sample was also relatively mild.

Second, the considerable heterogeneity in economic sentiment across Districts characterized by the wide shaded area in Figure 1 is consistent with research (Owyang, Piger, and Wall, 2005; Owyang, Rapach, and Wall, 2009) showing that states both differ in the levels of growth that they experience in recessions and expansion phases of the business cycle and that individual states and Districts are often out of sync with the national business cycle. However, the proportion of Districts that experience negative economic sentiment clearly rises during US recessions. Over the sample in Figure 1, of the 64 months in which the United States was in recession, according to the NBER, 27 of these were associated with sentiment's simultaneously being negative in all 12 Districts.

We also observe from Figure 1 that national economic sentiment does not always lie in the middle of the District-level estimates, as one might expect if national sentiment were the equal-weighted aggregation of regional sentiment. In fact, since 2021 national sentiment has been more positive than in most individual Districts. That is, the Beige Book narrative about the national economy as a whole has been more upbeat than in the average District, as represented in Figure 1 by consensus sentiment. These differences between consensus and national economic sentiment in Figure 1 are statistically significant.[8]

To test whether, when characterizing national economic sentiment, the Beige Book, intentionally or otherwise, pays more attention to sentiment in some Districts than others, we let the data speak, as it were: We estimate the combination weights on the 12 measures of District-level sentiment that best explain (under squared error loss) movements in national sentiment by regressing national sentiment on the 12 District-level sentiment indices. Table 1, column (1) shows that, on average over our sample from 1970 through 2024, national sentiment loads more heavily on sentiment in specific Districts.[9]

Figure 1: District-Level and National Economic Sentiment (Trailing 3-Month Rolling Average)
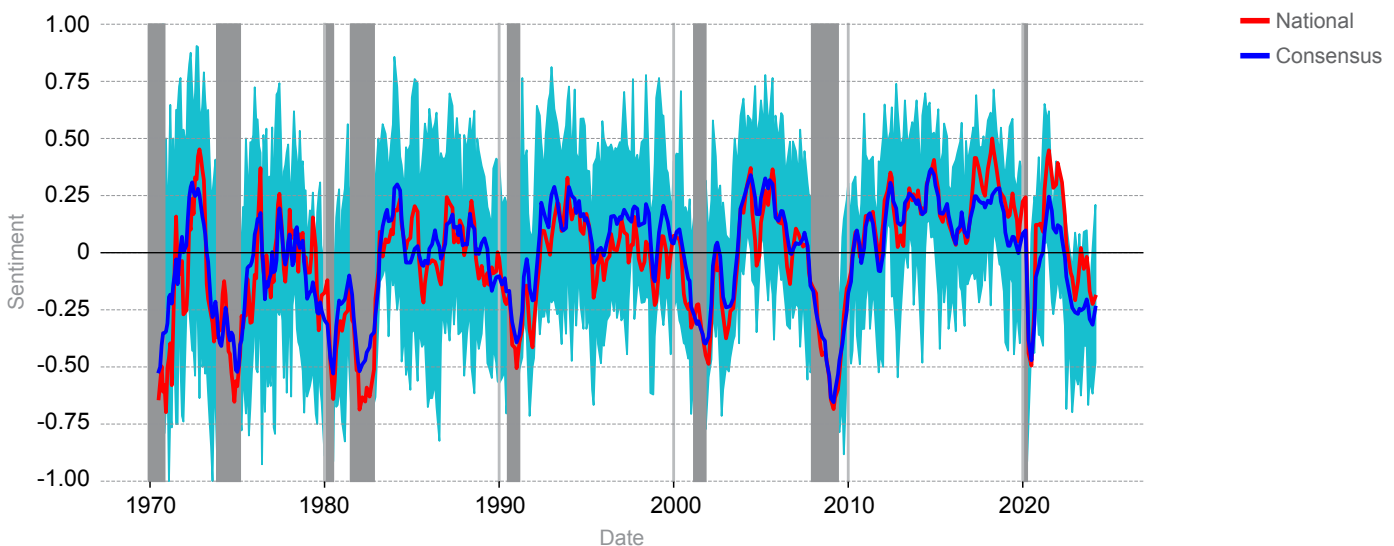
Note: Shaded area captures the upper and lower boundaries of the data ranges capturing the 12 measures of District-level sentiment extracted from each District report in the Beige Book. "Consensus" sentiment is the equal-weighted average of sentiment in each of the 12 Districts. Vertical shaded bars represent NBER recessions.

We might expect that the differential weights on the 12 Districts reflect the relative size of each District, if when writing the national narrative the authors implicitly or explicitly acknowledge a District's size. The difference between national and consensus sentiment, as seen in Figure 1, already suggests that Districts do not appear to be treated equally. As one measure of size, we look at the population of each District. The Federal Reserve Bank of San Francisco provides population estimates for the 12 Districts, discussing how the population shares have changed over time.[10] To pick a date reasonably close to the middle of our sample, in 2000 San Francisco was the largest Federal Reserve District, with a population share (relative to the US resident population) of 19 percent, followed by Atlanta with 15 percent, Chicago with 13 percent, Richmond with 10 percent, and Dallas with 8 percent. Supportive of the national narrative's reflecting the greater population size of these Districts, all five of these Districts receive positive and statistically significant weights in Table 1. In contrast, the smallest four Districts by population in 2000 were Cleveland, Kansas City, and St. Louis, all with 4 percent each, and Minneapolis with 3 percent of the US population. Of these four Districts, only Kansas City is statistically significant in Table 1, column (1). Of the midsized Districts—Boston with 5 percent, New York with 7 percent, and Philadelphia with 8 percent of the population share—Boston and New York are statistically significant, but Philadelphia is not. The fact that the estimated weights on Cleveland, Minneapolis, Philadelphia, and St. Louis are all statistically insignificant from zero, at the 1 percent, 5 percent, and 10 percent levels, indicates that, on average over our sample, these states are under-weighted in the national narrative in the Beige Book.

It is also noteworthy that the R-squared estimate from Table 1, column (1) indicates that 30 percent of the variation in national sentiment is *not* explained by the typical movements in District sentiment. National economic sentiment is more than a stable linear aggregation of District-level economic sentiment.

Predictive Power of District-Level Sentiment

To measure the information content of the economic sentiment indices in Figure 1 and the possibly differential informational content of national and District-level sentiment, we test their predictive ability for the US business cycle. We focus on analyzing the predictive power of both the national and District-level sentiment indices for US recessions as identified by the NBER on a monthly basis. The NBER maintains a chronology of US business cycles.[11] The chronology identifies the dates of peaks and troughs that define economic recessions and expansions. Recessions are defined retrospectively as broad-based declines in economic activity spread across the economy that last for more than a few months.

The forecasting problem that we consider is to predict, using data on each District's sentiment from the Beige Book released in month t, whether the US economy will be in a recession h Beige Book releases into the future, the timing of which we will refer to as simply month $t+h$. This does mean that we are abusing the definition of a "month" in these predictive regressions, since, strictly speaking, our *h-month*-ahead forecasts are *h-Beige Book*-ahead forecasts.[12] As emphasized by Armesto et al. (2009), the Beige Book does not have a regular release schedule. When relating to external economic data, this irregularity can raise questions about how to average Beige Book information to correspond with the period (say, monthly or quarterly) covered by the economic data release. In our application, given our use of monthly NBER business cycle data, we simply relate the Beige Book in a given month to the NBER data at that month, or leads and lags, and therefore neither control for when in the month the Beige Book was published nor for the unequal temporal gap between each of the eight Beige Books typically published per year.

We accordingly estimate h-month-ahead predictive (binary) regressions seeking to explain NBER-dated US expansions in month $t+h$ with District- and national-level sentiment data. Given that the NBER tends to publish its recession dates at a lag, often many months after the beginning of a recession, it remains of interest to backcast ($h<0$) expansions/recessions and to nowcast ($h=0$) and forecast ($h>0$) future ones. Indeed, it is an empirical question whether District-level economic sentiment, as quantified via FinBERT, is a forward-, coincident-, or backward-looking indicator of the US business cycle. The advantage of our in-sample tests of predictive power is that they benefit from longer samples. But, as is well known, there is no reason to expect any in-sample relationship to hold out of sample. This said, our sentiment measures do benefit from being computed in real time and so do not themselves suffer from look-ahead biases.

Table 1, columns (2) through (6) examine the predictive ability of the District-level and national sentiment indices for US business cycle phases (expansions and contractions) as captured by NBER-dated turning points. Looking at the (pseudo) R-squared metrics first that provide a measure of goodness of fit, we see that the predictive power of the sentiment indices is highest at $h=0$, namely as an indicator of the current business cycle phase. We repeat that, given that the NBER tends to date a recession up to a year after it has started, this result is consistent with the sentiment indices' still having meaningful predictive power.

Table 1: Explaining National Economic Sentiment with District-Level Sentiment and Forecasting NBER Business Cycle Phases

	(1) National	(2) NBER: h=-6	(3) NBER: h=-3	(4) NBER: h=0	(5) NBER: h=3	(6) NBER: h=6
National		1.91*	2.26	-0.03	-1.52	-0.60
		(1.75)	(1.63)	(-0.02)	(-1.23)	(-0.56)
Boston	0.11***	-0.72	0.16	1.30	3.44***	4.34***
	(2.81)	(-0.99)	(0.17)	(1.36)	(3.49)	(5.24)
Chicago	0.15***	1.97**	2.37**	2.53**	2.23**	1.75*
	(3.29)	(2.15)	(2.50)	(2.00)	(1.97)	(1.76)
Atlanta	0.07**	-0.89	0.28	1.20	1.67*	1.64*
	(2.09)	(-1.26)	(0.36)	(1.29)	(1.76)	(1.89)
Cleveland	0.06	-0.88	-0.64	-0.30	1.31	1.16
	(1.04)	(-1.19)	(-0.74)	(-0.29)	(1.36)	(1.43)
Dallas	0.08**	0.59	0.73	1.11	0.27	-0.17
	(2.35)	(0.96)	(0.94)	(1.46)	(0.36)	(-0.25)
Kansas City	0.08**	0.02	-0.08	-0.35	0.24	0.25
	(1.98)	(0.03)	(-0.12)	(-0.36)	(0.30)	(0.38)
Minneapolis	0.04	0.98	1.92**	2.04**	-0.31	-0.77
	(0.94)	(1.28)	(2.38)	(2.20)	(-0.45)	(-1.08)
New York	0.12***	0.63	-0.21	1.28	0.32	-0.16
	(3.84)	(0.97)	(-0.24)	(1.15)	(0.35)	(-0.15)
Philadelphia	0.05	1.15*	1.33**	1.79***	-0.01	-1.08*
	(1.52)	(1.90)	(2.07)	(2.67)	(-0.01)	(-1.82)
Richmond	0.12***	-1.31*	0.85	2.84**	1.56*	0.86
	(3.18)	(-1.78)	(0.91)	(2.31)	(1.73)	(1.13)
San Francisco	0.13***	2.93***	2.15**	2.92**	1.16	-0.96
	(3.77)	(3.88)	(2.26)	(2.41)	(1.23)	(-1.08)
St. Louis	0.03	-0.18	-0.84	-0.41	0.79	1.67**
	(1.19)	(-0.25)	(-1.09)	(-0.48)	(0.98)	(2.55)
Constant	-0.00	3.00***	3.86***	5.51***	4.63***	4.03***
	(-0.35)	(10.47)	(10.05)	(6.11)	(8.49)	(10.38)
R-squared	0.70	0.35	0.54	0.67	0.50	0.38
Observations	455	454	454	454	453	450
National = 0		0.08	0.10	0.98	0.22	0.57
Districts = 0		0.00	0.00	0.00	0.00	0.00

Column (1) reports the results from regressing national sentiment on sentiment in each of the 12 Districts. Columns (2) through (6) report the results from estimating (in-sample) a logit model relating NBER expansion dummies h-months-ahead to national and District-level economic sentiment at time t, as extracted from successive Beige Book issues from May 1970 through March 2024. t-statistics in parentheses; for column (1) these are calculated using HAC standard errors with 5 lags. *** (and bold red underline) indicates statistical significance at the 1 percent level, ** (and bold red) at the 5 percent level, and * (and red) at the 10 percent level. National=0 and Districts=0 are p-values testing the statistical significance of the national and 12 District-level sentiment indices, respectively.

While, as expected, we see this predictive power drop off as we look further into the future (or back into the past), sentiment retains statistically significant information useful for dating the business cycle. We emphasize that sentiment measures from the Beige Book can be constructed on a timely basis in real time given that, unlike traditional quantitative variables, such as GDP, the Beige Book is published more rapidly to feed into successive FOMC meetings.

We now turn to the estimated coefficients in the predictive regressions for NBER expansions seen in Table 1. These coefficients indicate the weight that should be attached to each District when constructing a forecast of the national business cycle. Positive coefficients indicate that as sentiment rises, there is a higher probability of being in an expansionary phase of the business cycle. Analysis of these coefficients also provides a means of testing whether District-level sentiment provides useful information over and above that contained in the national economic summary. The p-values reported at the bottom of Table 1 provide a direct test of whether national or District-level sentiment indicators contain statistically useful information for explaining US business cycle dynamics. Looking at these p-values, and the associated estimated coefficients, we see that national economic sentiment is not statistically significant at h=0 or when looking into the future (h=3 and h=6). In contrast,

our estimates of District-level economic sentiment offer value-added relative to national sentiment. The null hypothesis that District-level sentiment is uninformative is always rejected, with p-values of 0.00.[13] Inspecting the estimated coefficients on the different Federal Reserve Districts at h=0, we see that sentiment in Chicago, Minneapolis, Philadelphia, Richmond, and San Francisco is most helpful in predicting current US business cycle turning points. The sign on these estimated coefficients is always positive, indicating that as sentiment improves in each District, there is a higher chance of the US economy's being in an expansionary phase. Boston then becomes the most informative District both three- and six-months ahead, indicating that the most informative District(s) can vary by forecast horizon.[14] In summary, the observed heterogeneity in District-level economic sentiment can be used, over and above the information contained in national economic sentiment, to better forecast US recessions.

Finally, Figure 2 uses the logit model from Table 1, column (4) to plot the conditional probability that the US economy is in recession in the month captured by that Beige Book publication. We see that historically these model-based probabilities match up well with the months that the NBER subsequently defined as recessions. Higher probabilities in Figure 2 are associated with NBER-defined recessions: when the model predicts a recession probability of greater than 80 percent or so, there is a near certainty that the NBER subsequently declared the United States to have indeed been in a recession. From the mid-1980s, and the start of a period of economic stability often called the "Great Moderation" by economists, until the onset of the COVID-19 pandemic in March 2020, we also observe no instances of false alarms, that is, of recession probabilities' spiking above around 40 percent without a recession actually taking place. Since COVID-19 we have seen, similarly to pre-Great

Moderation, much more volatility in recession probabilities. As of March 2024, the probability of a national recession was low, conditioning on the national and District-level sentiment data as extracted from the March 2024 Beige Book. But this probability has bounced around considerably since the pandemic-induced recession of early 2020.

Conclusion

This *Economic Commentary* shows how natural language processing methods can be used to automate the quantification of text-based regional economic intelligence. Our analysis reveals meaningful economic differences across Federal Reserve Districts in their economic sentiment indices. It also documents that national economic sentiment, as measured by the Beige Book's own national economic summary, is not always the equal-weighted aggregation of District-level economic sentiment. Since the pandemic, national economic sentiment has tended to paint a rosier picture than that experienced in many individual Districts based on our model's processing of the textual data.

We validate the use of text-based estimates of regional economic sentiment from the Beige Book by showing that they contain statistically useful information about US business cycle phases, opening up their routine use in (real time) nowcasting and forecasting models. Importantly, District-level sentiment contains information useful for identifying US business cycle phases over and above the information contained in the Beige Book's national economic summary. We view these findings as supporting the attention that the regional Reserve Banks give to firsthand reports of economic activity in their respective Districts.

Figure 2. Probability That the United States Is in Recession Conditioning on National and District-Level Economic Sentiment
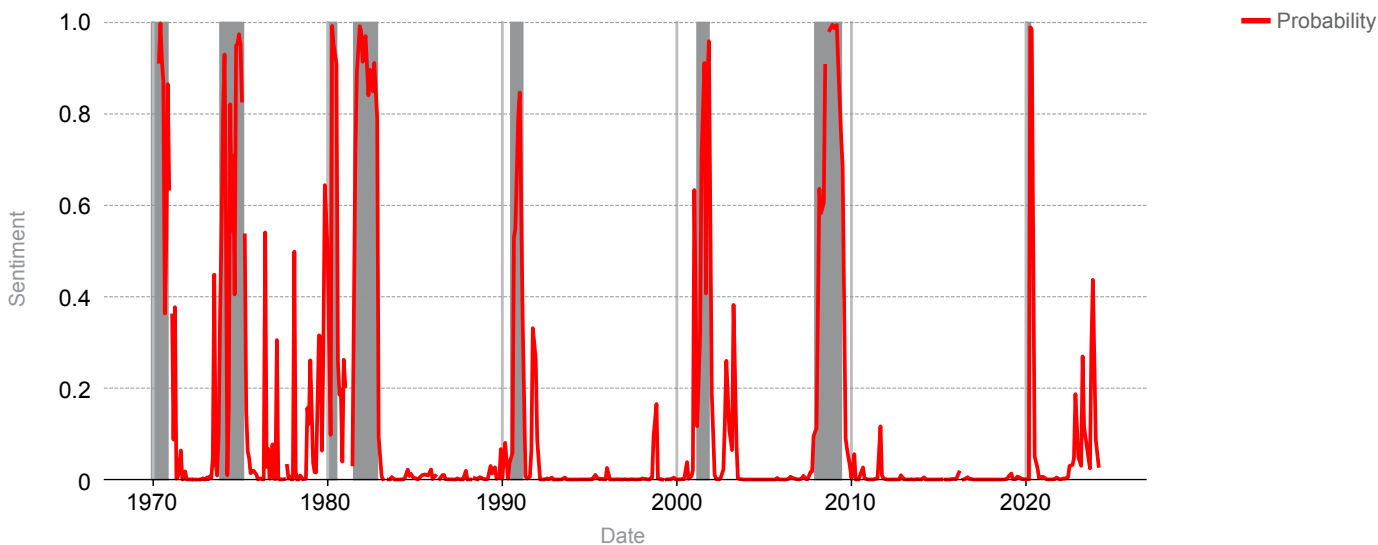
Note: The figure displays the predicted probability of recession from the estimated logit model in Table 1, column (4). Shaded areas represent NBER recessions.

Endnotes

1. Through much of the 1970s the Beige Book was published more than eight times per year. We exploit these additional reports in our analysis below.

2. For example, see Carroll, Fuhrer, and Wilcox (1994) and Souleles (2004).

3. Various methods from natural language processing have been used to convert textual (qualitative) data into quantitative data that can then be used in econometric models; see Gentzkow, Kelly, and Taddy (2019) and Ash and Hansen (2023) for an overview.

4. The financial corpus used to train FinBERT comprises the following documents: corporate annual and quarterly 10-K and 10-Q filings of Russell 3000 firms between 1994 and 2019, financial analyst reports issued for S&P 500 firms between 2003 and 2012 from the Thomson Investext database, and earnings conference call transcripts of 7,740 public firms between 2004 and 2019 from the SeekingAlpha website.

5. Such an analysis might use a different stream of the natural language processing literature, topic models (for example, see Larsen and Thorsrud, 2019), and decompose the Beige Book text into "news" topics associated with different aspects of the economy (real economic activity, employment, prices, financial markets, future outlook, and so on). We leave such an analysis for future research.

6. While we smooth the sentiment indices when presenting them in Figure 1, the rest of our analysis uses the raw, underlying sentiment estimates. Results are robust to this choice.

7. This finding complements and extends to the present day the results in Sadique et al. (2013). They also find that national sentiment, extracted from the Beige Book, has predictive power for business cycle turning points that differentiate expansions from contractions. We also note that national and consensus economic sentiment appears to be capturing something distinct from popular measures of overall economic activity such as the Federal Reserve Bank of Chicago's national activity index. The correlation between the Chicago Fed's index and the two sentiment measures seen in Figure 1 is low, at around 0.3.

8. A regression estimated with heteroscedasticity- and autocorrelation-consistent (HAC) standard errors of the difference between national and consensus economic sentiment on an intercept indicates that national sentiment was statistically below consensus sentiment prior to the pandemic (t-statistic of -3) but statistically greater than consensus sentiment since March 2020 (t-statistic of 6.3) even though the full-sample correlation between national and consensus sentiment in Figure 1 is solid at 0.83.

9. While our sample involves analyzing 468 Beige Book issues from May 1970 through March 2024, the number of time-series observations in our tables is slightly less than this, because for a handful of months the Beige Book is not available for all 12 Districts; for example, the January 1971 Federal Reserve Bank of Boston report is not available. We delete these occasional missing observations before estimating the regressions.

10. See https://www.frbsf.org/education/publications/doctor-econ/2001/may/federal-reserve-districts/.

11. See https://www.nber.org/research/business-cycle-dating.

12. With eight Beige Books typically published per year, we are, in fact, forecasting whether the United States is in recession in about $h(1+1/3)$ months' time.

13. This result complements those of Owyang, Piger, and Wall (2015), who find that state-level employment growth substantially improves nowcasts and very short-horizon forecasts of the US business cycle phase. That is, regional data can help when forecasting the national economy.

14. This is not to say, however, that sentiment in the other Districts is not helpful in explaining economic activity in that specific District. Armesto et al. (2009) find that the regional sections of the Beige Book are, in general, informative for District-level employment movements.

References

Angeletos, George-Marios, Fabrice Collard, and Harris Dellas. 2018. "Quantifying Confidence." *Econometrica* 86 (5): 1689–1726. https://doi.org/10.3982/ECTA13079.

Armesto, Michelle T., Rubén Hernández-Murillo, Michael T. Owyang, and Jeremy Piger. 2009. "Measuring the Information Content of the Beige Book: A Mixed Data Sampling Approach." *Journal of Money, Credit and Banking* 41 (1): 35–55. https://doi.org/10.1111/j.1538-4616.2008.00186.x.

Ash, Elliott, and Stephen Hansen. 2023. "Text Algorithms in Economics." *Annual Review of Economics* 15 (1): 659–88. https://doi.org/10.1146/annurev-economics-082222-074352.

Baker, Scott R., Nicholas Bloom, and Steven J. Davis. 2016. "Measuring Economic Policy Uncertainty." *The Quarterly Journal of Economics* 131 (4): 1593–1636. https://doi.org/10.1093/qje/qjw024.

Balke, Nathan S., and D'Ann Petersen. 2002. "How Well Does the Beige Book Reflect Economic Activity? Evaluating Qualitative Information Quantitatively." *Journal of Money, Credit, and Banking* 34 (1): 114–36. https://doi.org/10.1353/mcb.2002.0024.

Barsky, Robert B., and Eric R. Sims. 2012. "Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence." *American Economic Review* 102 (4): 1343–77. https://doi.org/10.1257/aer.102.4.1343.

Carroll, Christopher D., Jeffrey C. Fuhrer, and David W. Wilcox. 1994. "Does Consumer Sentiment Forecast Household Spending? If So, Why?" *American Economic Review* 84 (5): 1397–1408. https://www.jstor.org/stable/2117779.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. "BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding." Working paper arXiv:1810.04805. http://arxiv.org/abs/1810.04805.

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy. 2019. "Text as Data." *Journal of Economic Literature* 57 (3): 535–74. https://doi.org/10.1257/jel.20181020.

Gorodnichenko, Yuriy, Tho Pham, and Oleksandr Talavera. 2023. "The Voice of Monetary Policy." *American Economic Review* 113 (2): 548–84. https://doi.org/10.1257/aer.20220129.

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): 806–41. https://doi.org/10.1111/1911-3846.12832.

Larsen, Vegard H., and Leif A. Thorsrud. 2019. "The Value of News for Economic Developments." *Journal of Econometrics* 210 (1): 203–18. https://doi.org/10.1016/j.jeconom.2018.11.013.

Owyang, Michael T., Jeremy Piger, and Howard J. Wall. 2005. "Business Cycle Phases in U.S. States." *Review of Economics and Statistics* 87 (4): 604–16. https://doi.org/10.1162/003465305775098198.

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): 847–66. https://doi.org/10.1111/jmcb.12228.

Owyang, Michael T., David E. Rapach, and Howard J. Wall. 2009. "States and the Business Cycle." *Journal of Urban Economics* 65 (2): 181–94. https://doi.org/10.1016/j.jue.2008.11.001.

Sadique, Shibley, Francis In, Madhu Veeraraghavan, and Paul Wachtel. 2013. "Soft Information and Economic Activity: Evidence from the Beige Book." *Journal of Macroeconomics* 37 (September): 81–92. https://doi.org/10.1016/j.jmacro.2013.01.004.

Shiller, Robert J. 2019. *Narrative Economics: How Stories Go Viral and Drive Major Economic Events.* Princeton: Princeton University Press. https://doi.org/10.1515/9780691189970.

Souleles, Nicholas S. 2004. "Expectations, Heterogeneous Forecast Errors, and Consumption: Micro Evidence from the Michigan Consumer Sentiment Surveys." *Journal of Money, Credit, and Banking* 36 (1): 39–72. https://doi.org/10.1353/mcb.2004.0007.

Thorsrud, Leif A. 2020. "Words Are the New Numbers: A Newsy Coincident Index of the Business Cycle." *Journal of Business & Economic Statistics* 38 (2): 393–409. https://doi.org/10.1080/07350015.2018.1506344.

Zavodny, Madeline, and Donna K. Ginther. 2005. "Does the Beige Book Move Financial Markets?" *Southern Economic Journal* 72 (1): 138-151. https://doi.org/10.2307/20062098.