FEDERAL RESERVE BANK of ST. LOUIS

RESEARCH DIVISION

Working Paper Series

"Measuring the Information Content of the Beige Book: A Mixed Data Sampling

Michelle T. Armesto,
Rubén Hernández-Murillo,
Michael T. Owyang
and
Jeremy M. Piger

Working Paper 2007-010B
https://doi.org/10.20955/wp.2007.010

May 2008

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
P.O. Box 442
St. Louis, MO 63166

Measuring the Information Content of the Beige Book: A Mixed Data Sampling Approach*

Michelle T. Armesto†

Rubén Hernández-Murillo†

Michael T. Owyang‡†

and

Jeremy Piger§

keywords: data sampling frequency, textual analysis, DICTION, Beige Book

May 8, 2008

Abstract

Studies of the predictive ability of the Federal Reserve's Beige Book, an anecdotal measure of regional economic conditions, for aggregate output and employment have proven inconclusive. This might be attributed, in part, to the irregular release schedule of the Beige Book. In this paper, we use a model that allows for data sampling at mixed frequencies to analyze the predictive power of the Beige Book for both aggregate and regional data. We find that the Beige Book's national summary and District reports predict GDP and aggregate employment and that most District reports provide information content for regional employment. In addition, there appears to be an asymmetry in the predictive content of the Beige Book language.
[JEL: C50, E27, R11]

*The authors thank Oscar Jorda, Kevin Kliesen, Ivana Komunjer, Barbara Rossi, and Enrique Sentana for comments and suggestions. We thank Nathan Balke and D'Ann Petersen for the use of their Beige Book measures. Heidi Beyer-Powe, Kristie M. Engemann, George Essig, Christopher Martinek, and Deborah Roisman provided research assistance. The views expressed in this paper are the authors' alone and do not reflect the views of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

†Research Division. Federal Reserve Bank of St. Louis. P.O. Box 442. St. Louis, MO 63166-0442.

‡Corresponding author: owyang@stls.frb.org

§Department of Economics. 1285 University of Oregon. Eugene, OR 97403.

1 Introduction

The charter of the Federal Reserve System established twelve geographic Districts with an aim, in part, to provide local influence in monetary policy and to exploit regional information advantages in evaluating aggregate conditions. The latter objective is fulfilled through the collection and dissemination of the *Beige Book*, an anecdotal summary of regional economic conditions. How well the Beige Book reflects both regional and aggregate conditions is a question that has been studied in a number of papers (e.g., Fettig et al. (1999); Balke and Petersen (2002); Ginther and Zavodny (2001)). More recently, some studies have also analyzed how well the Beige Book predicts financial variables such as interest rates and equity prices (e.g., Zavodny and Ginther (2005)). Unfortunately, only a weak consensus has been reached on the Beige Book as an indicator of either current or future aggregate economic conditions. Various studies of the value of the Beige Book as a predictor of aggregate conditions have found it to depend on (i) the section of the Beige Book (e.g., the average of regional conditions or the national summary) used as a regressor, (ii) the aggregate measure of economic conditions (e.g., gross domestic product (GDP), employment) being evaluated, and (iii) the timing of the Beige Book release relative to the aggregate data release.

Balke and Petersen (2002), for example, studied Beige Book reports released between July 1983 and January 1997, giving numerical scores to the national summary, national sectoral reports, and each Federal Reserve District report. Because the Beige Book is not released at regular intervals, Balke and Petersen took particular care in specifying the timing of the Beige Book release dates relative to other indicators of economic activity. They found that their Beige Book measures tracked current real GDP growth well. They also concluded that the Beige Book had predictive content for current- and next-quarter real GDP growth beyond that of alternative indicators such as the Blue Chip consensus forecast. Balke and Petersen additionally concluded that the Beige Book appeared to identify turning points sooner than most of the alternative indicators. These results contrast with previous studies (e.g., Fettig et al. (1999)), which found that the Beige Book explained a minority of the variation in current-quarter real GDP growth.[1]

Other papers have examined the information content of the regional Beige Book reports to

[1] In their analysis, Fettig et al. (1999) included the period between 1970 and mid-1983 when the Beige Book (known then as the Red Book) was confidential and was not released to the public. Balke and Petersen (2002) argued that differences in their findings may be attributable to the timing of the private sector forecasts with which the Beige Book is compared.

determine their ability to track current regional economic activity. Balke and Yücel (2000), for example, examined whether the Dallas Fed's Beige Book tracks current Texas real gross state product (GSP) growth and employment growth. Using the Beige Book scores of Balke and Petersen (2002) for the Eleventh Federal Reserve District and quarterly estimates of Texas real GSP growth (Berger and Phillips (1995)), Balke and Yücel found that the regional Beige Book indices track the Texas economy well. They also found that the Beige Book contains some information not present in other indicators of economic activity. Ginther and Zavodny (2001) performed a similar exercise using the Atlanta Fed's Beige Book and measures of quarterly regional economic activity in the Sixth District.[2] They found that the Sixth District Beige Book score tracks the regional economy well. However, because the regional and national Beige Book scores are highly correlated, Ginther and Zavodny (2001) found that the regional Beige Book provides little additional information when the national Beige Book summary is included in the analysis.

This paper addresses three issues in the current literature. First, because the information contained in the Beige Book is, for the most part, informal and anecdotal in nature, the Beige Book narrative on economic activity is difficult to quantify. The language of individual Districts varies greatly.[3] Previous studies that relied on human readers required careful and precise efforts to minimize the subjective scoring of the Beige Book. In Balke and Petersen (2002), for example, after removing references to calendar years, each author read the reports in random sequence; the resulting index was constructed from an average of the two authors' scores. Thus, even when keeping the same readers, these efforts make replication and extension of the dataset virtually impossible. Our solution to this problem is to utilize linguistics software that evaluates the degree of "optimism" and "pessimism" of each Beige Book release (Hart, 2000*b*, 2001). While the primary advantage of this type of software is replicability conditional on a known dictionary, it also removes a large degree of the subjectivity associated with human readers.

[2]Ginther and Zavodny (2001) defined Sixth District–level measures of activity summing employment, income, and quarterly estimates of GSP across the six states that comprise the District (Alabama, Florida, Georgia, and parts of Louisiana, Mississippi, and Tennessee). They used growth rates in their analysis and expressed employment in per capita terms to control for population growth. Ginther and Zavodny (2001) also used the Balke and Petersen (2002) scores. However, Ginther and Zavodny matched each Beige Book to a particular quarter, averaging all of the Beige Book scores associated with that quarter.

[3]Some Districts describe the *pace* of economic activity (e.g., "Atlanta reported a robust pace," national summary April 2005). Other Districts describe changes in the *level* of economic activity (e.g., "Business activity in the Eighth District continued to expand," St. Louis April 2005). The descriptions of growth rely heavily on the use of adjectives such as *solid*, *moderate*, or *modest*, and even *better* (as in, "They expect the pace of hiring this spring to be better than last year," Philadelphia April 2005).

Second, we formally address the issue of mixing sampling frequencies. Since the Beige Book is released irregularly, the econometrician is left with the problem of either averaging Beige Book scores, attributing multiple Beige Books to each GDP release, or (arbitrarily) assigning a specific Beige Book score to each GDP release. To solve this problem, we adopt an econometric specification (dubbed the MIDAS model in Ghysels et al. (2004)) designed to jointly handle mixed data-sampling frequencies and control the number of estimated parameters. This framework allows us to incorporate the information content from multiple Beige Books by weighting each Beige Book score as a function of the elapsed time between Beige Book and data releases. Thus, the model evaluates the effect of a sequence of Beige Books leading up to an economic data release.

Finally, we evaluate the predictive power of each region's Beige Book using District-level employment data. The intent of the Beige Book is to provide a measure of local conditions under the assumption that these local conditions reflect, at least in part, current or future aggregate conditions. However, if the Beige Book fails to reflect even regional conditions, the likelihood that it might predict national conditions should be virtually zero. While a few of the papers discussed above do consider the regional sections of the Beige Book, none of these studies analyzes each of the twelve Fed Districts in a consistent framework.

The remainder of this paper is organized as follows. In section 2 we outline the empirical model and describe the estimation technique. We discuss the data and summarize the content of the Beige Book, the timing of releases, and the method used to score the Beige Book in section 3. We then address the construction of the regional economic indicator (District-level employment) used for the disaggregate regressions that follow. In section 4 we present the results of the estimation using the growth rates in real GDP and aggregate employment. Here, we provide some extensions of the baseline model, including expanding the set of regressors. We present the results from the estimation of District-level employment growth on the relevant section of the Beige Book in section 5.

2 Econometric Model

Estimating a time-series model in which a measure of the Beige Book enters as a right-hand-side variable requires modeling data sampled at different frequencies. Measures of output (e.g., GDP) are often sampled quarterly, while employment and industrial production are sampled monthly. The Beige Book, on the other hand, is released at irregular intervals, meaning that varying numbers of Beige Books can be observed in any quarterly GDP cycle and multiple employment reports may be released between Beige Books. This difference in sampling frequency is more than an econometric complication; Balke and Petersen (2002) found that the timing of the Beige Book release is an important factor in determining the Beige Book's predictive power. This suggests a decline in the information content of the Beige Book as a function of increased elapsed time between the Beige Book and data releases.

Unlike other studies in which data are simply time-averaged, we can exploit this variation in frequencies by employing the MIDAS model suggested in Ghysels et al. (2004). MIDAS addresses the difference in sampling frequencies between the left- and right-hand-side variables by employing a weighted time aggregation. In this application, the weights are chosen to be functions of the elapsed time between sampled data and an estimated vector of hyper-parameters.

Let the growth rate of the economic condition of interest (e.g., employment or GDP) and the Beige Book measure be denoted Y and X, respectively. We use the notation Y_t to indicate that variable Y is sampled once between period $t-1$ and period t (e.g., once per month in the case of employment and once per quarter in the case of GDP). In contrast, we use the notation $X_t^{(m)}$ to indicate that variable X is sampled more frequently, m times in the same period (in our case, weekly). Let k denote the elapsed time (in weeks) between the Beige Book and the end of the data-reporting period, and let $k^{\max}$ denote an exogenously set time limit in weeks, such that for $k > k^{\max}$, the Beige Book has a negligible predictive effect on Y_t.

The model can then be written in the following form:

$$Y_t = \alpha + \beta \sum_{k=0}^{k^{\max}} \Gamma(k, \boldsymbol{\theta}) X_{t-k/m}^{(m)} + \varepsilon_t, \tag{1}$$

where α is an intercept term, β is a slope coefficient, and $\varepsilon_t \backsim i.i.d.\left(0, \sigma^2\right)$. The lag in $X_{t-k/m}^{(m)}$ is expressed as a fraction of the unit interval between $t-1$ and t. That is, if t indicates the monthly frequency and the Beige Book is assumed to be sampled weekly, then $m = 4$. If, for example, the Beige Book is released two weeks before the end of month t, $X_{t-2/4}^{(4)}$ represents the Beige Book measure as of its release date. Since there is no other Beige Book release that month, we set

$X^{(4)}_{t-k/4} = 0$ for $k = 1, 3, 4$ to indicate that there is no new information in those dates.

The weights $\Gamma(k, \boldsymbol{\theta})$ are governed both by the elapsed time k and by an n-dimensional vector of hyper-parameters $\boldsymbol{\theta}$.[4] We impose a two-parameter Almon function (that is, $n = 2$) for the weights:

$$\Gamma(k, \boldsymbol{\theta}) = \frac{\exp\left(\theta^1 k + \theta^2 k^2\right)}{\sum_{k=0}^{k^{\max}} \exp\left(\theta^1 k + \theta^2 k^2\right)}, \tag{2}$$

where $\sum_{k=0}^{k^{\max}} \Gamma(k, \boldsymbol{\theta})$ is normalized to 1. Essentially, (2) imposes a functional dependence of the model coefficients on the time elapsed between the release of the left- and right-hand-side variables that reduces the dimensionality of the econometric model (1).

As we previously noted, the effect of a series of Beige Books in the framework outlined in (1) is cumulative up to $k^{\max}$ weeks prior to the economic data release.[5] The MIDAS model used here stands in contrast with the model of Balke and Petersen (2002), who assigned a quarterly GDP growth rate to each Beige Book, essentially attaching multiple Beige Book scores to a single economic data release. MIDAS views the events of the quarter in totality as affecting the final economic outcome.

Our model of the information content of the Beige Book includes two measures of economic conditions as described in the text of the reports: optimism and pessimism. We can modify the baseline MIDAS model in the following manner:

$$Y_t = \alpha + \beta_o \sum_{k=0}^{k^{\max}} \Gamma_o(k, \boldsymbol{\theta}_o) X^{(m)}_{o,t-k/m} + \beta_p \sum_{k=0}^{k^{\max}} \Gamma_p(k, \boldsymbol{\theta}_p) X^{(m)}_{p,t-k/m} + \varepsilon_t, \tag{3}$$

where the subscripts on Γ_i, β_i, and $X^{(m)}_{i,t-k/m}$, $i \in \{o, p\}$, represent optimism and pessimism, respectively. Equation (3) then represents an econometric model with a vector of parameters $\Psi = \left[\alpha, \beta_o, \boldsymbol{\theta}'_o, \beta_p, \boldsymbol{\theta}'_p\right]'$ and an innovation variance σ^2 that can be estimated via nonlinear least squares.

Before proceeding to the empirical results, it is important to consider the manner in which parameter values from the estimated model can be interpreted. In particular, we are interested in distinguishing the effects of the coefficients on the MIDAS terms, the β_i's, from the hyper-

[4] Choice of the functional form of $\Gamma(k, \theta)$ and the dimensionality of θ is discussed at length in Ghysels et al. (2007).

[5] For simplicity, in the present model, we have chosen the maximum number of lags, $k^{\max} = 24$ weeks, exogenously. In general, the data will determine $k^{\max}$ as long as $\theta^1 > 0$ and $\theta^2 < 0$.

parameters in the weighting functions, the θ_i's. The former indicate the effect of the Beige Book measure on the variable of interest. The relevant test for whether the Beige Book has predictive power is a test against the null hypothesis of $\beta_i = 0$. On the other hand, the latter do not reflect the information content of the Beige Book, per se, but test which Beige Books are relevant for predicting Y_t. As a special case, the weighting function $\Gamma_i(k, \boldsymbol{\theta}_i)$ reduces to simple averaging (as in the model used in Ginther and Zavodny (2001)) when both parameters θ_i^1 and θ_i^2 are zero. Thus, a joint test of both hyper-parameters equaling zero, $\theta_i^1 = \theta_i^2 = 0$, is not a test of the lack of predictive power for the Beige Book but is a test of equality of weight for past Beige Books back through $k^{\max}$.

3 Data

As was alluded to above, any study of the Beige Book's information content requires a measure of the Beige Book's outlook. In this section, we provide a brief discussion of the Beige Book itself. We then outline the technique used to generate a measure of the Beige Book's outlook. Finally, we summarize the economic conditions data used in the Federal Reserve District-level regressions.

3.1 Beige Book data

In this paper, we consider the Beige Book reports from May 1970 through July 2005. The Beige Book, officially called "Summary of Commentary on Current Economic Conditions," contains a description of economic activity in each of the twelve Federal Reserve Districts and one national report based on the Districts' reports.[6] Each District report is three pages long, although in recent years, there have been some exceptions. Each report includes a summary paragraph of economic conditions in the District and a more detailed description in labeled categories, which include descriptions of individual industries, labor market conditions, and in some cases commentary on prices and wages. The national summary also includes an introductory paragraph and detailed commentary on individual economic indicators.[7]

[6] The Beige Book is available online at <http://www.federalreserve.gov/FOMC/BeigeBook/2008/>. It is released eight times per year, roughly two weeks before each scheduled meeting of the Federal Open Market Committee (FOMC). Historic Beige Book reports are available from the Federal Reserve Bank of Minneapolis Beige Book archive <http://www.mpls.frb.org/bb/>.

[7] Balke and Petersen (2002) provide a more detailed description of the Beige Book. See also <http://www.fmcenter.org/site/pp.asp?c=8fLGJTOyHpE\&b=224810> for a description of the Beige Book methodology.

3.2 Textual Analysis

Since the Beige Book is anecdotal, we must construct a measure of the general tenor of the Beige Book language. While previous studies used subjective measures (panels of graders, etc.), we employ an alternative measure of Beige Book optimism and pessimism constructed by linguistics programs.[8] With the use of the textual-analysis software DICTION 5.0, our study generates systematic measures of the levels of optimism and pessimism contained in the narrative of the Beige Book.[9] The program can process an unlimited number of texts using a corpus of several thousand words, generating scores for five general features: certainty, activity, optimism, realism, and commonality. These scores are computed using scores from thirty-five subcategories. In our case, for example, the optimism score is computed by adding the standardized word frequencies of the subcategories labeled as *optimism increasing* by DICTION (praise, satisfaction, and inspiration), while our pessimism score is computed by adding the standardized word frequencies of the subcategories labeled *optimism decreasing* by DICTION (blame, hardship, and denial). The software is customizable with dictionaries (word lists) chosen for specific research purposes.[10]

The use of DICTION has several advantages over human coding. First, textual-analysis techniques based on pre-existing search rules and algorithms are systematic and, thus, free from criticisms of research subjectivity and bias that might be levied against human coding. Second, given identical dictionaries, the resulting analysis is perfectly reproducible, and new observations can be added to the sample without fear of influencing the scoring process.[11] The main disadvantage of textual analysis is that, although the program counts words characterized as optimistic or pessimistic based on linguistic theory (Hart, 1984, 1987, 2000*a*,*b*, 2001), it is incapable of providing analysis of language conditional on the context of the particular statement.

[8] In contrast to previous studies, we do not attempt to quantify directly the strength of regional economic growth; instead, we quantify the qualitative features of the Beige Book reports.

[9] DICTION 5.0 (Hart, 2000*b*, 2001) is described by its seller as "a scientific method for determining the tone of a verbal message." DICTION is a dictionary-based content analysis program that contains the types of words most frequently used in contemporary American public discourse (Hart, 1984). DICTION has been employed extensively to analyze contemporary discourse including speeches of politicians (Hart, 1984, 2000*a*,*b*; Bligh et al., 2003, 2004; Hart and Jarvis, 1997); speeches of Federal Reserve policymakers (Bligh and Hess, 2007, 2005); annual reports to stockholders (Yuthas et al., 2002); and other business communications, including earnings press releases (Davis et al., 2006).

[10] The Praise score, for example, includes words that isolate social qualities (*witty*), physical qualities (*strong*), intellectual qualities (*reasonable*), entrepreneurial qualities (*successful*), and moral qualities (*good*). See the DICTION 5.0 manual for more details, `http://www.dictionsoftware.com/files/dictionmanual.pdf`.

[11] The dictionaries used to construct our Beige Book measure are available in Diction 5.0.

The textual-analysis software outputs two series of scores for each Beige Book report: optimism and pessimism. Table 1 presents the summary statistics of the Beige Book optimism and pessimism scores for the aggregate measures and each of the regional Reserve Banks. Because the scaling of our textual measure differs from Balke and Petersen's human-coded measure used in previous studies, a direct comparison is impossible. Figure 1, however, plots both our optimism and pessimism scores for the Beige Book's national summary alongside the human-coded measures for the period in which the latter is available. While there are some differences, the two measures appear sufficiently correlated to provide some assurance that the DICTION-based and the human-coded measures are indeed capturing similar characteristics.[12]

3.3 Measures of Economic Conditions

We consider two measures of aggregate economic conditions: the seasonally adjusted annualized growth rates of aggregate employment and final GDP. Employment is released monthly, while final GDP is released quarterly. Our disaggregate measure of economic conditions is Federal Reserve District employment growth for the period February 1972 to December 2005. State-level employment cannot be used as a regional measure because some states are shared by multiple Federal Reserve Districts. Unfortunately, no output measure is available at both a disaggregated level and a reasonable frequency. Gross state product, the closest approximation to GDP at the state level, is annual and, until recently, available only with a two-year lag.

Following Balke and Petersen (2002), we evaluate the Beige Book's predictive ability for the most current data on employment and GDP (measured as of December 2006), rather than using the *vintage* employment and GDP data released in real time. Consistent with the model given in 1, the value of k is measured as the number of weeks from the release date of the Beige Book to the end of the sampling period for employment or GDP (that is the last day of the quarter for GDP or the month for employment). Alternatively, we could measure k as the distance between the Beige Book release date and the release date of a particular vintage of GDP and employment. The reason we do not measure k in this way is because we are interested in whether Beige Books

[12]The correlation of the human-coded measure with our optimism score is 0.13, and the correlation with our pessimism score is -0.28. Over Balke and Petersen's sample period (July 1983 to January 1997) the correlation between our optimism and pessimism scores is 0.01; however, the correlation over the entire sample period (May 1970 to November 2005) is about -0.24.

(released during the quarter in question) have predictive content for GDP and not vice versa. For example, the release date of the so-called *final* estimate of GDP usually occurs two months into the next quarter following the reporting period, and would thus fall later than Beige Books released after the end of the GDP reporting period but before the *final* GDP release.[13]

We approximate Federal Reserve District employment data by aggregating monthly metropolitan total nonfarm payroll employment compiled from the Current Employment Statistics survey.[14] We aggregated the series by Federal Reserve District as follows: We first aggregated employment from the metro areas (defined as of December 2003) that were entirely contained in each Fed District. If the metro area crossed District borders, we assigned total nonfarm employment to the Fed District containing the metro area's central business District.[15] Finally, if the central business District also crossed Federal Reserve District borders, we assigned the metro area's total nonfarm employment to both Districts.[16,17] Table 2 presents summary statistics for the monthly employment growth series both for the nation and the twelve regional Federal Reserve Bank Districts, along with summary statistics for GDP growth.

4 National Results

The goal of this study is to establish a set of stylized facts regarding the information content of the Beige Book. While the Beige Book is primarily composed of anecdotal evidence of regional economic conditions, it does contain a section summarizing national conditions. Moreover, the average of the District conditions should reflect, to some extent, the underlying state of the aggregate economy.

[13]This timing scheme has the same flavor as Balke and Petersen's. They tie GDP releases to Beige Books released in the same quarter, although their methodology prevents them from including lagged Beige Books.

[14]For the Twelfth District, Seattle, Tacoma, and Spokane were omitted due to missing data.

[15]There were 16 metro areas in this case: Albuquerque, NM; Alexandria, LA; Allentown-Bethlehem-Easton, PA-NJ; Augusta-Richmond County, GA-SC; Bloomington, IN; Cincinnati-Middletown, OH-KY-IN; Clarksville, TN-KY; Columbia, MO; Fayetteville-Springdale-Rogers, AR-MO; Kingsport-Bristol, TN-VA; New York-Newark-Edison, NY-NJ-PA; Omaha-Council Bluffs, NE-IA; Parkersburg-Marietta, WV-OH; Philadelphia-Camden-Wilmington, PA-NJ-DE-MD; Sioux City, IA-NE-SD; and Terre Haute, IN.

[16]There were only three metro areas in this case: Fort Smith, AR-OK; Texarkana, TX-AR; and Huntington-Ashland, WV-KY-OH.

[17]The employment level series for each metropolitan area were seasonally adjusted using the Census Bureau's X-12 ARIMA program. The level series were then aggregated by Fed District. Finally, the growth series were computed from the District aggregates and filtered to eliminate outliers.

4.1 Baseline Results

The first two columns of Table 3 report the baseline regression results (based on equation 3) for GDP on two Beige Book measures – the national summary and the average of the District scores. We find that the national summary and the District average scores of the Beige Book, respectively, explain about 15 percent and 18 percent of the variation in GDP growth.[18] Note that, regardless of which measure of the Beige Book is employed, both the optimism and pessimism scores are statistically significant. As expected, optimism (pessimism) enters with a positive (negative) sign, indicating that more optimistic (pessimistic) Beige Books are associated with higher (lower) overall GDP growth. Similar conclusions can be drawn from the Beige Book's predictive content for aggregate employment in columns three and four of Table 3.[19]

It is possible that the information content of the Beige Book language has evolved over time, as the reports became public information and the potential for receiving more media attention than in the past increased. We conducted a simple of test of subsample stability running the regressions for the aggregate indicators with an exogenously chosen break around the middle of our sample in January 1989. We also ran the regressions for the front half of the sample (from 1972 to 1988). We found that the results for the post-1989 subsample were qualitatively consistent with our reported results for the average of the District scores. The MIDAS coefficients were, in addition, somewhat larger in magnitude for the front half of the sample.[20] However, the full sample is short and we did not pursue additional stability tests.

Table 3 also includes tests against alternative models in which the hyper-parameters θ_i^1 and θ_i^2 in the weighting function are zero for either $i = o$ or $i = p$. We cannot reject that past Beige Books' optimism scores enter the GDP regression with equal weights. As mentioned before, these results do not indicate that the Beige Book optimism is uncorrelated with GDP. Instead, the inability to reject that the θs are jointly zero suggests that past Beige Book optimism enters as a simple average over the past $(k^{max} + 1)/m$ weeks. In contrast, we reject equal weighting of past pessimism scores.

[18]In both cases, the explanatory power of the Beige Book that we find is less than that reported by Balke and Petersen (2002). These figures, however, are not directly comparable because (i) the measure of the Beige Book is different, (ii) the dataset in this paper is longer, and (iii) the model used here combines multiple Beige Books into a single measure.

[19]The hyper-parameters are unidentified under the null hypothesis of $\beta_i = 0$. Reported levels of significance assume a standard chi-squared distribution as an approximation to the true distribution.

[20]These results are available upon request.

The estimated hyper-parameters suggest that more pessimistic Beige Books closer to the end of the quarter should be weighted more in predicting that quarter's GDP growth. The third and fourth columns of Table 3 show that we reject that optimism (when using the national summary and the District average) and pessimism (when using the national summary) enter the regression for aggregate employment as a simple average.

A typical set of weights is illustrated in Figure 2. The weights in this chart correspond to the optimism $\boldsymbol{\theta} = (0.1444, -0.0082)$ in the MIDAS regression for aggregate employment using the national summary of the Beige Book. Note that the peak does not occur for contemporaneous releases of the Beige Book. In this case, the peak of the weighting function for optimism appears at about nine weeks. This result appears to contrast with the findings of Balke and Petersen (2002)—that the Beige Books closer to the release date of the economic indicator contain more relevant information.[21]

4.2 Robustness

The preceding results confirm the predictive content of the Beige Book; in this subsection, we verify the veracity of these conclusions in the presence of additional information. For example, other studies have shown that the predictive power of the Beige Book diminishes once, say, lagged dependent variables are added. In our econometric framework, incorporating additional explanatory variables is easily accomplished by including additional terms in equation (3).[22] We can also consider predicting the left-hand-side variable over different horizons, replacing Y_{t+h} for Y_t, with $h > 0$.

In this section, we conduct the following in-sample forecasting exercise. First, we consider a forecasting horizon h and estimate the following autoregressive model for the aggregate economic indicator Y_t:

$$Y_{t+h} = \alpha + \phi(L)Y_{t-1} + \varepsilon_t, \tag{4}$$

[21] Running the exercises reported in Table 2 of Balke and Petersen (2002) with our optimism and pessimism measures indicates that the average of the District measures are statistically significant when explaining current-quarter GDP growth, whereas the national summary optimism and pessimism are significant for explaining next-quarter GDP growth. These results are available from the authors upon request.

[22] Previous studies have also considered the information content of the Beige Book in the presence of economic forecasts, such as the Blue Chip Economic Indicators. The addition of economic forecasts to our econometric specification, however, requires an assumption on timing. While we could, in principle, use judgment in comparing the timing of the forecasts, they may still overlap with the information in previous Beige Books. This confounds our ability to disentangle the Beige Book's predictive ability from economic forecasts.

where

$$\phi(L) = (\phi_1 + \phi_2 L + \phi_3 L^3 + \cdots + \phi_q L^{q-1}).$$

With fixed h, we then choose the lag order q minimizing the Bayesian information criterion (considering a maximum value for the lag order equal to 24). Once the lag order q is chosen, we augment the model with the MIDAS term for the Beige Book scores:

$$Y_{t+h} = \alpha + \phi(L)Y_{t-1} + \beta_o \sum_{k=0}^{k^{\max}} \Gamma_o(k, \boldsymbol{\theta}_o) X^{(m)}_{ot-k/m} + \beta_p \sum_{k=0}^{k^{\max}} \Gamma_p(k, \boldsymbol{\theta}_p) X^{(m)}_{pt-k/m} + \varepsilon_t. \tag{5}$$

Conditional on h and q, the parameters $\Psi = \left[\alpha, \beta_o, \boldsymbol{\theta}'_o, \beta_p, \boldsymbol{\theta}'_p\right]'$, along with the coefficients of the lag polynomial $\phi(L)$, are re-estimated.

Table 4 summarizes the results of this exercise for GDP and aggregate employment for different forecasting horizons h. The upper panel reveals that the Beige Book language retains predictive information for GDP for short horizons up to three quarters ahead ($h = 3$) when using either the national summary or the average of the District scores. Only the optimism measure is significant for $h = 1$ or $h = 2$, but both coefficients are again significant and of the expected sign for $h = 3$. For horizons of four quarters, the optimism coefficient is no longer statistically significant; and, although significant, the pessimism coefficient reverses sign. In the case of employment growth, the lower panel of Table 4 illustrates that the Beige Book language retains predictive information even after the inclusion of lagged employment for horizons up to twelve months when using the average of the District measures, although only the optimism coefficient is statistically significant for horizons of one month or longer. The optimism coefficient is also significant for a horizon of twelve months. Using the national summary yields statistically significant optimism coefficients for $h = 0$ and $h = 12$. The score for pessimism is also statistically significant at a twelve-month horizon (although with a positive sign).

5 District-Level Results

While verifying that the Beige Book has some predictive power for aggregate economic conditions is reassuring, another primary purpose of the Beige Book lies in its reflection of regional conditions.

Other studies have examined the extent to which particular Districts' Beige Book sections reflect the economic conditions in their areas. Here, we broaden those studies to determine which Districts, if any, produce Beige Book sections correlated with their employment conditions.

5.1 Baseline Results

Table 5 summarizes the results for each of the regional Federal Reserve Banks using (metropolitan) District payroll employment as the measure of economic activity. Note that seven of the twelve Federal Reserve Districts exhibit statistically significant coefficients for *both* Beige Book optimism and pessimism scores. Additionally, in the New York and Cleveland Fed Districts, only pessimism is statistically significant; and, in the Minneapolis and Kansas City Fed Districts, only optimism is statistically significant.[23] In the Atlanta Fed District, however, neither optimism nor pessimism has significant explanatory power for its District's employment. In the cases in which the coefficients are significant, they appear with the expected sign.

Our results suggest two regional asymmetries in assessing the predictive power of the Beige Book's optimism and pessimism scores. First, some regional Feds' Beige Book sections more accurately reflect the true (developing) nature of their Districts. The explained variation of regional employment varies from about 1.6 percent in the St. Louis Fed District to about 18 percent in the Boston Fed District.[24] When both optimism and pessimism scores are statistically significant, the explained variation in metropolitan employment growth averages about 8 percent but can be as low as 1.6 percent. Second, the level of conservatism in the language of the Beige Book varies across Federal Reserve Districts. Five Districts (Boston, Richmond, Minneapolis, Kansas City, and Dallas) have substantially greater-than-average coefficients on their optimism scores, implying that even small increases (decreases) in their optimism are correlated with disproportionately large increases (decreases) in their District's employment growth. This result suggests that the language of these Districts tends toward conservatism, where substantial changes in their District fundamentals produce only small variation in the tenor of their wording. For pessimistic language, Cleveland, Richmond, and Dallas have the greatest (negative) correlation with declining regional

[23]One possible explanation for this result lies in the construction of the District employment series. For Kansas City and Minneapolis, agriculture may play a large role in the general tenor of the Beige Book. Since rural employment is omitted from our measure, fluctuations in economic conditions stemming from agriculture may be unaccounted for.

[24]The Beige Book explains only about 0.4 percent of the Atlanta Fed's regional employment. However, neither optimism nor pessimism is statistically significant in that District.

employment. Again, this suggests that the Richmond and Dallas Districts' section of the Beige Book are some of the most conservatively worded.

Asymmetry across Federal Reserve Districts is not limited to variation in the coefficients in equation (3). Columns four and five of Table 5 consider the null hypothesis that the weighting function hyper-parameters are jointly zero for at least one of the Beige Book scores. We find that this null hypothesis is rejected for four of the twelve Federal Reserve Districts: New York, Minneapolis, Kansas City, and San Francisco. In all of these Districts, the hypothesis of equal weights is rejected for only one of the scores. These results differ from the aggregate employment results in that neither optimism nor pessimism (when using the national summary) was found to enter as simple averages and only optimism (when using the average of the District measures) was found to enter as a simple average.[25]

5.2 Robustness

As with the aggregate, we can consider whether the Beige Book District reports possess predictive content in excess of the inclusion of additional information (in this case, lags of District employment growth) and over different (in-sample) forecasting horizons, according to the model outlined in equations (4) and (5). Table 6 presents the results of these tests. Consistent with the findings for aggregate employment in section 4.2, the addition of lagged employment mitigates the information content of the Beige Book for District-level employment, as shown in the panel corresponding to $h = 0$. In this case, fewer Districts exhibit statistically significant coefficients for both optimism and pessimism scores and three (in addition to Atlanta) lose statistical significance for both parameters. Nevertheless, when considering extended forecasting horizons of up to twelve months ahead, it is clear that the Beige Book language of many Districts continues to exhibit predictive content from at least one of the Beige Book scores. As we consider longer horizons, however, the signs of pessimism (and the sign of optimism for the New York District) tend to reverse.[26]

The notion that the regional sections of the Beige Book are, at least in part, informative for

[25] Although applying the MIDAS methodology to the human-coded measure of Balke and Petersen (2002) results in statistically significant estimates for both the aggregate and District measures, a comparison of our results with theirs would not be a simple task because their Beige Book scores do not allow for the asymmetries between different language components of the Beige Book text. Some results in this direction are available from the authors upon request.

[26] Although not shown in Table 6, adding lags of employment to the regressions reduces the absolute value of both the optimism and pessimism coefficients.

District-level employment fluctuations is encouraging. That the Beige Book remains informative when the regression includes lagged employment is further evidence that even anecdotal information gathered at the regional level can be a useful indicator of the general tenor of the economy. However, still troubling is the inability of some Districts' sections to explain their own regional economies. In the previous discussion, we alluded to some possible explanations as to why the Beige Book might be less informative in some Districts. In particular, which industries (e.g., agriculture, mining) are the focus of a District's Beige Book might affect its performance in these regressions. Another possibility is that District-wide MSA employment is not representative of the sample drawn on by the regional Bank when forming its conclusions about regional conditions. Two scenarios arise. First, information in the Beige Book emphasizes conditions in non-metropolitan (e.g., rural) employment. Second, information in the Beige Book is concentrated on the Federal Reserve Bank Branch cities. We consider the latter scenario in the next subsection.

5.3 Branch-Level Analysis

One consideration in interpreting the results in the preceding subsections is the manner in which each District acquires the anecdotes that comprise the Beige Book. In particular, some Districts may overemphasize information originating from metropolitan areas that have branches. To test for this possibility, we re-estimate (3) using the main office and branch city employment growth as the dependent variable. The results of these branch city regressions are presented in Table 7. This table also includes log-likelihood ratio tests for the alternatives of simple averages in the weights function. While results vary substantially across Districts, some important insights can be ascertained from this set of branch-level regressions.

The salient result is that, for the vast majority of the Districts, the predictive acumen of the Beige Book is enhanced when attention is restricted to the main office city. In virtually every case, at least one of the optimism and pessimism coefficients is statistically significant, and in more than half of the Districts, both coefficients are statistically significant for the main office city (denoted in italics in Table 7). Furthermore, the magnitude of the statistically significant coefficient rises when considering the main office city versus the entire District (in Table 5), especially for the pessimism coefficient. This result, while not surprising, confirms that the Beige Book reflects more closely the economic performance of the main office city than that of the District. We do not, however, interpret this as a point of emphasis of each District's Beige Book report. Instead, information from the main office city may simply be more readily available or of higher quality.[27]

A second feature of these city-level results is the Beige Book's lack of predictive power for a number of branch cities. For 8 of the 35 branch cities, their District's Beige Book exhibits no predictive power. For an additional 13 of these cities, only a single term – either optimism or pessimism – is statistically significant.

These results, however, fail to shed any additional light on the inability of the Atlanta District Beige Book to predict economic activity in its region. Our results indicate that restricting attention to branch cities only weakly enhances the performance of the Beige Book in informing fluctuations in the Atlanta District. Our results show that only Birmingham and Jacksonville exhibit statistical significance. These results are seemingly in contrast to Ginther and Zavodny (2001), who find that the Atlanta District Beige Book does predict regional economic activity. Three explanations are possible: (1) The DICTION-based measure of the Beige Book is inappropriate for the Atlanta District, which could relate back to the contextual issue we discussed above. (2) Employment is a poor measure of economic activity in the District. (3) Their data omitted some of the major hurricanes that struck the region in the early part of the decade.

6 Conclusion

Previous studies of the information content of the Beige Book have been unable to model the timing differential between the sampling frequency of the data and that of the Beige Book itself. In this paper, we have shown that this timing issue is important, at least at some levels of disaggregation. Our model not only confirms the predictive power of the Beige Book, but also provides a sense of the asymmetry underlying the language of the Beige Book, both at the aggregate and regional levels. At the aggregate level, the asymmetry occurs in the nature of the information provided by the optimistic and pessimistic language of the Beige Book's national summary. Our results suggest that information from recent pessimism scores should receive more weight in predicting GDP, while optimism scores can be weighted equally.

[27] Another alternative is that the Beige Book reports are *anchored* to the main branch city. Campbell and Sharpe (2007), for example, argue that forecasters tend to anchor toward past forecasts, creating a systematic bias. We leave explicit tests of these alternatives to future research.

At the regional level, asymmetries exist in both the value and type of information relating to District-level employment. While the Beige Book can have predictive value for most Federal Reserve Districts' employment growth, the information for each region is conveyed in different linguistic components. In some regions, pessimism is the key component relating the Beige Book to District employment. In other regions, optimism – or both characteristics – reflects the state of the economy. Moreover, how accurately the Beige Book predicts employment depends on the District in question. In many instances, the Beige Book is not a reflection of the region as a whole but, instead, more highly correlated with the main office city where information may be of higher quality and more easily obtained.

One might naturally wonder how these findings can be useful outside the Federal Reserve. Hernández-Murillo and Owyang (2006) found that forecasts of national employment can be improved by aggregating regional employment in a spatial model. Their disaggregate forecasts were constructed using lags of monthly employment growth rates. If the Beige Book is indeed informative beyond lagged regional employment, the addition of anecdotal information could improve even aggregate forecasting accuracy.

References

Balke, Nathan S. and D'Ann Petersen, "How Well Does the Beige Book Reflect Economic Activity? Evaluating Qualitative Information Quantitatively," *Journal of Money, Credit, and Banking*, February 2002, *34* (1), pp. 114–136.

Balke, Nathan S. and Mine K. Yücel, "Evaluating the Eleventh District's Beige Book," *Federal Reserve Bank of Dallas Economic and Financial Review*, Fourth Quarter 2000, pp. 2–10.

Berger, Franklin D. and Keith R. Phillips, "A New Quarterly Output Measure for Texas," *Federal Reserve Bank of Dallas Economic Review*, Third Quarter 1995, pp. 16–23.

Bligh, Michelle C. and Gregory D. Hess, "A Quantitative Assessment of the Qualitative Aspects of Chairman Greenspan's Communication," Working paper, Claremont Graduate University and Claremont McKenna College, September 2005.

—, "The Power of Leading Subtly: Alan Greenspan, Rhetorical Leadership, and Monetary Policy," *The Leadership Quarterly*, April 2007, *18* (2), pp. 87–104.

Bligh, Michelle C., Jeffrey C. Kohles and James R. Meindl, "Textual Analysis of Leadership During Crisis: A Methodological Illustration of Responses to 9/11," in D. H. Nagao, ed., *Proceedings of the Sixty-third Annual Meeting of the Acdemy of Management (CD)*, 2003 .

—, "Charisma under Crisis: Presidential Leadership, Rhetoric, and Media Responses Before and After the September 11th Terrorist Attacks," *The Leadership Quarterly*, April 2004, *15* (2), pp. 211–239.

Campbell, Sean D. and Steven A. Sharpe, "Anchoring Bias in Consensus Forecasts and its Effect on Market Prices," Finance and Economics Discussion Series 2007-12, Federal Reserve Board, February 2007.

Davis, Angela K., Jeremy M. Piger and Lisa M. Sedor, "Beyond the Numbers: An Analysis of Optimistic and Pessimistic Language in Earnings Press Releases," Working paper 2006-005A, Federal Reserve Bank of St. Louis, January 2006.

Fettig, David, Arthur J. Rolnick and David E. Runkle, "The Federal Reserve's Beige Book: A Better Mirror Than Crystal Ball," *The Region*, March 1999, *13*(1), pp. 10–13,28–32.

Ghysels, Eric, Pedro Santa-Clara and Rossen Valkanov, "The MIDAS Touch: Mixed Data Sampling Regression Models," June 2004.

Ghysels, Eric, Arthur Sinko and Rossen Valkanov, "MIDAS Regressions: Further Results and New Directions," *Econometric Reviews*, 2007, *26*(1), pp. 53–90.

Ginther, Donna K. and Madeline Zavodny, "The Beige Book: Timely Information on the Regional Economy," *Federal Reserve Bank of Atlanta Economic Review*, Third Quarter 2001, *86*(3), pp. 19–29.

Hart, Roderick P., *Verbal Style and the Presidency: A Computer-Based Analysis*, Academic Press, Inc., Orlando, FL, 1984.

—, *The Sound of Leadership: Presidential Communication in the Modern Age*, University of Chicago Press, Chicago, IL, 1987.

—, *Campaign Talk: Why Elections are Good for Us*, Princeton University Press, Princeton, NJ, 2000*a*.

—, *DICTION 5.0: The Text-Analysis Program*, Scolari/Sage Publications, Thousand Oaks, CA, 2000*b*.

—, "Redeveloping DICTION: Theoretical Considerations," in Malk D West, ed., *Theory, Method, and Practice in Computer Content Analysis*, Ablex, New York, 2001 pp. 43–60.

Hart, Roderick P. and Sharon E. Jarvis, "Political Debate: Forms, Style, and Media," *American Behavioral Scientist*, August 1997, *40*(8), pp. 1095–1122.

Hernández-Murillo, Rubén and Michael T. Owyang, "The Information Content of Regional Employment Data for Forecasting Aggregate Conditions," *Economics Letters*, March 2006, *90*(3), pp. 335–339.

Yuthas, Kristi, Rodney Rogers and Jesse F. Dillard, "Communicative Action and Corporate Annual Reports," *Journal of Business Ethics*, November 2002, *41*(1-2), pp. 141–157.

Zavodny, Madeline and Donna K. Ginther, "Does the Beige Book Move Financial Markets?" *Southern Economic Journal*, July 2005, *72*(1), pp. 198–151.

Table 1: Summary Statistics: Beige Book Scores

	Optimism						Pessimism					
Score	n	mean	med	sd	min	max	n	mean	med	sd	min	max
Boston	295	1.68	1.60	0.61	0.00	4.00	295	1.90	1.80	0.97	0.00	4.70
New York	295	1.58	1.50	0.55	0.40	3.40	295	1.24	1.10	0.65	0.10	3.70
Philadelphia	295	1.59	1.60	0.58	0.20	3.30	295	0.92	0.90	0.52	0.00	3.50
Cleveland	295	1.53	1.40	0.75	0.30	4.40	295	1.32	1.20	0.57	0.10	3.60
Richmond	295	1.52	1.50	0.62	0.10	3.30	295	1.03	1.00	0.56	0.00	2.90
Atlanta	295	1.73	1.70	0.61	0.00	3.90	295	1.21	1.10	0.62	0.00	4.30
Chicago	295	1.68	1.60	0.61	0.00	3.60	295	1.41	1.30	0.71	0.00	3.80
St. Louis	295	1.40	1.40	0.58	0.00	2.90	295	0.99	0.90	0.54	0.00	3.60
Minneapolis	295	1.82	1.80	0.68	0.20	3.90	295	1.40	1.40	0.73	0.00	4.90
Kansas City	295	1.63	1.50	0.57	0.50	3.60	295	1.37	1.30	0.49	0.40	3.40
Dallas	295	1.40	1.40	0.61	0.10	3.00	295	1.19	1.10	0.50	0.10	3.00
San Francisco	295	1.92	1.90	0.68	0.30	3.90	295	1.39	1.40	0.65	0.00	3.80
Nat. Summary	319	1.90	1.80	0.62	0.40	3.60	319	1.29	1.20	0.63	0.10	3.60
Dist. Average	319	1.63	1.62	0.25	1.01	2.36	319	1.29	1.27	0.39	0.46	2.56

Sample for District Scores: 09/15/1971 to 12/01/2004

Sample for Nat. Summary and Dist. Average Scores: 05/20/1970 to 11/30/2005

Table 2: Summary Statistics: Employment and GDP Growth

	n	mean	med	sd	min	max
Boston	395	0.37	0.44	1.11	-3.15	4.88
New York	395	0.19	0.26	0.93	-3.05	3.49
Philadelphia	395	0.40	0.38	1.01	-3.66	3.74
Cleveland	395	0.34	0.35	1.09	-3.24	3.78
Richmond	395	0.75	0.75	1.03	-2.99	4.37
Atlanta	395	0.95	0.98	1.06	-2.77	4.36
Chicago	395	0.44	0.52	1.11	-3.45	3.53
St. Louis	395	0.56	0.62	1.29	-3.85	5.10
Minneapolis	395	0.79	0.76	1.18	-2.28	4.71
Kansas City	395	0.71	0.72	1.01	-3.15	3.45
Dallas	395	0.98	0.98	1.12	-2.76	4.75
San Francisco	395	0.82	0.91	1.02	-2.59	4.50
Agg. Employment	428	1.82	2.04	2.57	-8.85	15.89
GDP	143	3.17	3.30	3.45	-7.83	16.72

Growth is calculated as annualized period percent change.

Sample for District Employment: 1972:Feb to 2004:Dec

Sample for Agg. Employment: 1970:May to 2005:Dec

Sample for GPD: 1970:Q2 to 2005:Q4

Table 3: GDP and Aggregate Employment: MIDAS Regressions

	GDP		Employment	
	[1] NAT	[2] DAV	[3] NAT	[4] DAV
Const.	1.2908	-1.2434	-0.5355	$\mathbf{-2.1675}$
β_o	**11.9484**	**24.8990**	**10.4775**	**20.5524**
β_p	$\mathbf{-8.9024}$	$\mathbf{-11.7414}$	$\mathbf{-4.6644}$	$\mathbf{-7.8246}$
χ^2 statistic on $\theta_o^1 = \theta_o^2 = 0$	3.69	2.83	**7.40**	**21.30**
χ^2 statistic on $\theta_p^1 = \theta_p^2 = 0$	**8.16**	**6.44**	**8.91**	4.03
R^2	0.1488	0.1777	0.1485	0.2084
N. Obs.	143	143	428	428

Dependent variables are annual growth rates in GDP and employment.

Boldface indicates that the coefficient is statistically significant at 10% level.

β_o is the coefficient on optimism and β_p is the coefficient on pessimism from equation (3).

(θ_s^1, θ_s^2) for $s = o, p$ are the hyper-parameters of the weights function $\Gamma(k, \theta)$ defined in equation (2). The coefficients on the constraints are χ^2 statistics for likelihood ratio tests. The constraints tested imply constant weights.

NAT is the Beige Book score from the National Beige Book summary.

DAV is the average of scores from the District reports.

Table 4: In-Sample Forecasting Power of the Beige Book for GDP and Aggregate Employment

GDP

Score	$h=0$				$h=1$				$h=2$				$h=3$				$h=4$			
	q^*	β_o	β_p	R^2	q^*	β_o	β_p	R^2	q^*	β_o	β_p	R^2	q^*	β_o	β_p	R^2	q^*	β_o	β_p	R^2
NAT	1	▲	▼	0.16	1	▲	△	0.07	1	▲	▽	0.05	1	▲	▼	0.08	1	△	▲	0.07
DAV	1	▲	▼	0.18	1	▲	▽	0.10	1	▲	▽	0.10	1	▲	▼	0.09	1	△	▲	0.07

Aggregate Employment

Score	$h=0$				$h=1$				$h=3$				$h=6$				$h=12$			
	q^*	β_o	β_p	R^2	q^*	β_o	β_p	R^2	q^*	β_o	β_p	R^2	q^*	β_o	β_p	R^2	q^*	β_o	β_p	R^2
NAT	3	▲	▽	0.46	2	△	△	0.42	2	△	△	0.28	2	△	△	0.15	1	▲	▲	0.06
DAV	3	▲	▼	0.47	2	▲	▽	0.43	2	▲	▽	0.31	2	▲	▽	0.18	1	▲	△	0.07

NAT is the Beige Book score from the National Beige Book summary; DAV is the average of scores from the District reports.

For given h, q^* minimizes the $BIC(q)$ of the model $Y_{t+h} = \alpha + \phi(L)Y_{t-1} + \varepsilon_t$, where $\phi(L) = (\phi_1 + \phi_2 L + \phi_3 L^2 + \cdots + \phi_q L^{q-1})$, where Y_t is annual growth rates in the aggregate payroll employment series.

Fixing the lag order to q^*, the following augmented model is then estimated:

$Y_{t+h} = \alpha + \phi(L)Y_{t-1} + \beta_o \sum_{k=0}^{k^{\max}} \Gamma_o(k, \boldsymbol{\theta}_o) X_{o,t-k/m}^{(m)} + \beta_p \sum_{k=0}^{k^{\max}} \Gamma_p(k, \boldsymbol{\theta}_p) X_{p,t-k/m}^{(m)} + \varepsilon_t,$

where β_o is the coefficient on optimism, and β_p is the coefficient on pessimism.

The symbols △ and ▽ indicate a positive and a negative coefficient, respectively. Additionally, the symbols ▲ and ▼ indicate statistical significance at the 10% level.

Table 5: District Employment: MIDAS Regressions Baseline Model

District	α	β_o	β_p	$\theta_o^1 = \theta_o^2 = 0$	$\theta_p^1 = \theta_p^2 = 0$	R^2	No.Obs.
Boston	**−0.8684**	**6.1303**	**−1.5307**	2.44	0.46	0.1785	395
New York	**0.4057**	0.5203	**−1.6951**	1.51	**7.87**	0.0457	395
Philadelphia	−0.2442	**2.9381**	**−0.9548**	0.32	2.08	0.0425	395
Cleveland	**0.5752**	0.8783	**−2.1280**	0.58	1.70	0.0472	395
Richmond	**0.2794**	**3.8015**	**−2.8999**	3.98	2.61	0.1124	395
Atlanta	**1.1622**	−0.3910	−0.4826	0.49	0.83	0.0044	395
Chicago	0.2743	**1.4249**	**−1.0408**	1.15	3.78	0.0375	395
St. Louis	**0.5327**	**1.3570**	**−1.7614**	0.69	0.00	0.0158	395
Minneapolis	**−0.3499**	**3.7809**	−0.1013	**6.83**	0.00	0.1356	395
Kansas City	−0.1369	**3.7860**	−0.8434	**6.54**	1.51	0.1036	395
Dallas	0.3027	**4.9212**	**−2.4357**	3.61	0.77	0.1344	395
San Francisco	0.1060	**2.4601**	**−0.4518**	4.13	**5.09**	0.0721	395

Dependent variable is annual growth rates in district employment constructed by aggregating metropolitan employment.
Boldface indicates that the coefficient is statistically significant at 10% level.
β_o is the coefficient on optimism and β_p is the coefficient on pessimism from equation (3).
(θ_s^1, θ_s^2) for $s = o, p$ are the hyper-parameters of the weights function $\Gamma(k, \theta)$
defined in equation (2). The coefficients on the constraints are χ^2 statistics
for likelihood ratio tests. The constraints tested imply constant weights.

Table 6: In-Sample Forecasting Power of the Beige Book for District Employment

District	$h=0$				$h=1$				$h=3$				$h=6$				$h=12$			
	q^*	β_o	β_p	R^2	q^*	β_o	β_p	R^2	q^*	β_o	β_p	R^2	q^*	β_o	β_p	R^2	q^*	β_o	β_p	R^2
Boston	3	▲	▽	0.41	3	▲	▽	0.41	3	▲	▼	0.32	3	▲	▽	0.24	3	▲	▽	0.12
New York	6	▲	▼	0.32	5	▼	▲	0.30	3	▼	△	0.21	3	▼	△	0.13	3	▼	▲	0.08
Philadelphia	4	▲	▽	0.17	3	▲	▽	0.17	3	▲	△	0.15	3	▲	△	0.08	2	△	△	0.03
Cleveland	3	△	▽	0.28	3	△	△	0.26	3	△	△	0.17	2	△	△	0.08	1	△	△	0.02
Richmond	3	▲	▼	0.37	2	▲	▽	0.35	3	▲	▲	0.26	2	△	△	0.13	2	△	▲	0.07
Atlanta	3	▽	△	0.40	2	△	▲	0.39	2	△	△	0.23	1	△	△	0.10	1	△	▲	0.09
Chicago	3	▲	△	0.33	2	▲	△	0.30	3	▲	△	0.20	3	△	△	0.13	1	▽	▽	0.03
St. Louis	3	△	▽	0.18	2	▲	▽	0.18	2	▲	▽	0.08	1	▲	▼	0.07	1	△	▽	0.01
Minneapolis	4	▲	▽	0.27	3	▲	△	0.27	3	▲	▽	0.21	3	▲	△	0.16	1	▲	△	0.05
Kansas City	3	△	△	0.35	2	△	△	0.35	3	△	△	0.21	3	▲	△	0.15	2	▲	▽	0.12
Dallas	3	▲	▽	0.49	3	▲	▽	0.46	3	▲	▼	0.36	1	▲	▼	0.18	1	▲	△	0.04
San Francisco	3	▲	▼	0.52	2	△	▲	0.45	3	▽	▲	0.33	3	△	▲	0.22	2	▽	▲	0.12

For given h, q^* minimizes the $BIC(q)$ of the model $Y_{t+h} = \alpha + \phi(L)Y_{t-1} + \varepsilon_t$, where $\phi(L) = (\phi_1 + \phi_2 L + \phi_3 L^2 + \cdots + \phi_q L^{q-1})$, where Y_t is annual growth rates in the employment series constructed by aggregating metropolitan payroll employment in each Federal Reserve district.
Fixing the lag order to q^*, the following augmented model is then estimated:
$Y_{t+h} = \alpha + \phi(L)Y_{t-1} + \beta_o \sum_{k=0}^{k^{\max}} \Gamma_o(k, \boldsymbol{\theta}_o) X_{o,t-k/m}^{(m)} + \beta_p \sum_{k=0}^{k^{\max}} \Gamma_p(k, \boldsymbol{\theta}_p) X_{p,t-k/m}^{(m)} + \varepsilon_t$,
where β_o is the coefficient on optimism, and β_p is the coefficient on pessimism.
The symbols △ and ▽ indicate a positive and a negative coefficient, respectively. Additionally, the symbols ▲ and ▼ indicate statistical significance at the 10% level.

Table 7: District Employment: MIDAS Regressions Branch Cities

District/Office	α	β_o	β_p	$\theta_o^1=\theta_o^2=0$	$\theta_p^1=\theta_p^2=0$	R^2	No.Obs.
Boston							
Boston MA-NH PMSA	**−0.7724**	**5.6756**	**−1.5250**	1.19	0.00	0.1356	395
New York							
Buffalo-Niagara Falls NY	−0.0885	**3.2844**	**−3.1825**	3.40	0.24	0.0415	395
New York NY PMSA	**0.4585**	0.6419	**−2.7416**	2.89	**8.43**	0.0710	395
Philadelphia							
Philadelphia PA-NJ PMSA	0.0471	**1.9963**	**−1.5014**	2.71	0.72	0.0285	395
Cleveland							
Cincinnati OH-KY-IN PMSA	**0.7781**	0.5364	**−1.5957**	0.37	3.12	0.0193	395
Cleveland-Lorain-Elyria OH PMSA	0.3998	**1.0775**	**−2.2042**	1.92	1.03	0.0363	395
Pittsburgh PA	0.3056	**1.8116**	**−2.8409**	1.15	0.00	0.0661	395
Richmond							
Baltimore MD PMSA	0.0263	**4.2723**	**−4.0162**	1.77	0.20	0.0770	395
Charlotte-Gastonia-Rk Hill NC-SC	**0.5035**	**3.8955**	**−3.6091**	0.14	3.46	0.0600	395
Richmond-Petersburg VA	0.3157	**3.8029**	**−3.4180**	1.07	0.00	0.0731	395
Atlanta							
Atlanta GA	**1.4399**	−0.4874	−0.8381	0.33	1.50	0.0060	395
Birmingham AL	**1.0522**	−0.4701	**−1.4930**	0.09	0.18	0.0097	395
Jacksonville FL	**1.9141**	**−2.0487**	**−1.6766**	0.10	0.00	0.0211	395
Miami FL PMSA	**1.1498**	−1.2403	−1.0983	2.74	1.49	0.0122	395
New Orleans LA	**0.4667**	−0.1930	−0.5178	1.20	1.76	0.0062	395
Nashville TN	**0.7811**	1.7767	−1.6822	0.55	0.00	0.0083	395
Chicago							
Chicago IL PMSA	0.3167	**1.1925**	**−1.4245**	2.14	0.98	0.0423	395
Detroit MI PMSA	−0.3078	2.4957	−0.4674	0.80	3.59	0.0214	395
Oklahoma City OK	−0.3741	**2.8883**	0.8174	0.84	1.19	0.0256	395
Omaha NE-IA	**0.8130**	0.4006	**−1.0430**	0.00	1.43	0.0090	395
St. Louis							
Louisville KY-IN	0.4051	2.0131	−2.4643	0.33	0.02	0.0119	395
Little Rock-N Little Rock AR	0.3271	1.2036	0.9308	0.20	0.73	0.0073	395
Memphis TN-AR-MS	**0.9405**	0.4893	**−2.2686**	0.42	0.04	0.0112	395
St Louis MO-IL	0.2329	**2.2253**	**−1.8535**	2.99	2.41	0.0249	395
Minneapolis							
Minneapolis-St Paul MN-WI	**−0.3561**	**3.9691**	−0.3961	**5.84**	0.59	0.1322	395
Kansas City							
Denver CO PMSA	−0.1723	**5.2917**	**−1.5781**	**6.41**	2.73	0.1040	395
Kansas City MO-KS	−0.0798	**3.6301**	−1.5935	3.65	1.38	0.0329	395
Omaha NE-IA	0.4062	1.5516	−0.6153	0.12	0.38	0.0057	395
Dallas							
Dallas TX PMSA	**1.0438**	**5.3661**	**−6.2406**	1.94	0.14	0.1600	395
El Paso TX	**1.6645**	0.6100	**−4.9754**	0.55	0.12	0.0264	395
Houston TX PMSA	−0.3731	**6.7324**	−0.8994	**5.29**	3.31	0.1267	395
San Antonio TX	**1.0192**	**2.6434**	**−3.2910**	1.60	2.88	0.0471	395
San Francisco							
LA-Long Beach CA PMSA	**−0.5110**	**2.1949**	0.6931	2.58	1.75	0.0408	395
San Francisco CA PMSA	**−0.7705**	**3.0431**	0.6263	3.45	0.70	0.0435	395
Salt Lake City-Ogden UT	**0.5853**	**1.9050**	−0.8319	**5.51**	3.40	0.0359	395

Dependent variable is annual growth rates in metropolitan payroll employment.
Boldface indicates that the coefficient is statistically significant at 10% level.
β_o is the coefficient on optimism and β_p is the coefficient on pessimism from equation (3).
(θ_s^1, θ_s^2) for $s = o, p$ are the hyper-parameters of the weights function $\Gamma(k, \theta)$
defined in equation (2). The coefficients on the constraints are χ^2 statistics
for likelihood ratio tests. The constraints tested imply constant weights.

Figure 1: National Summary Scores

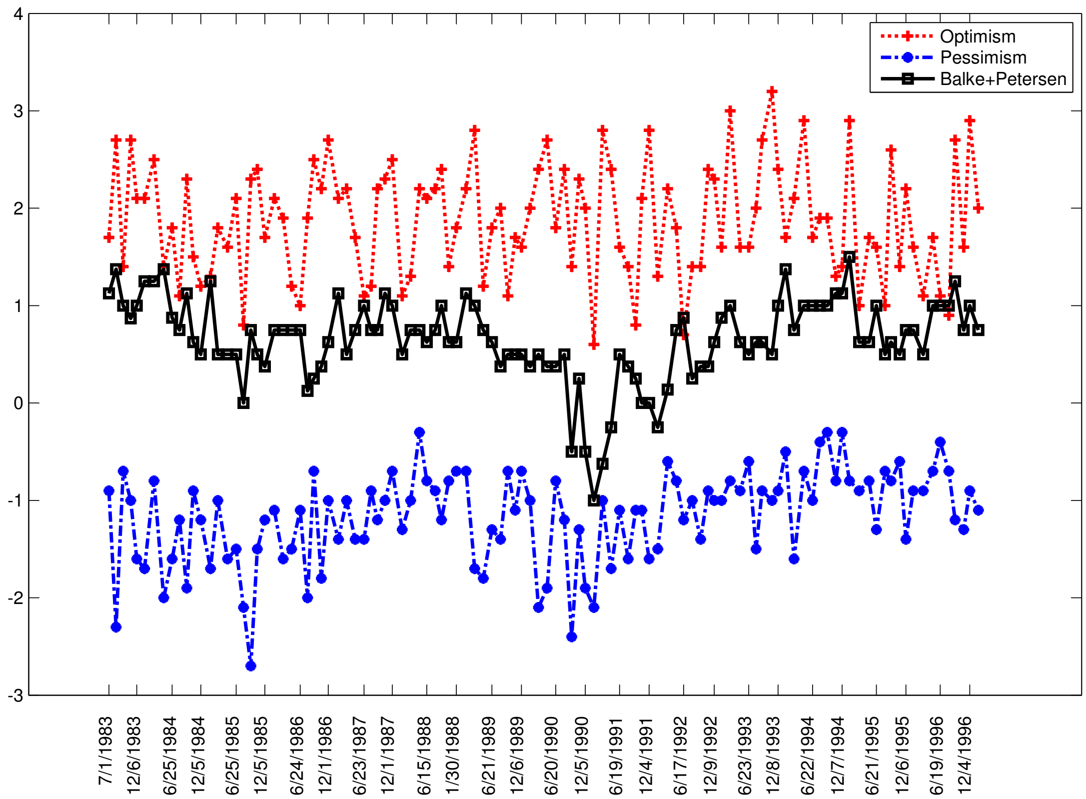
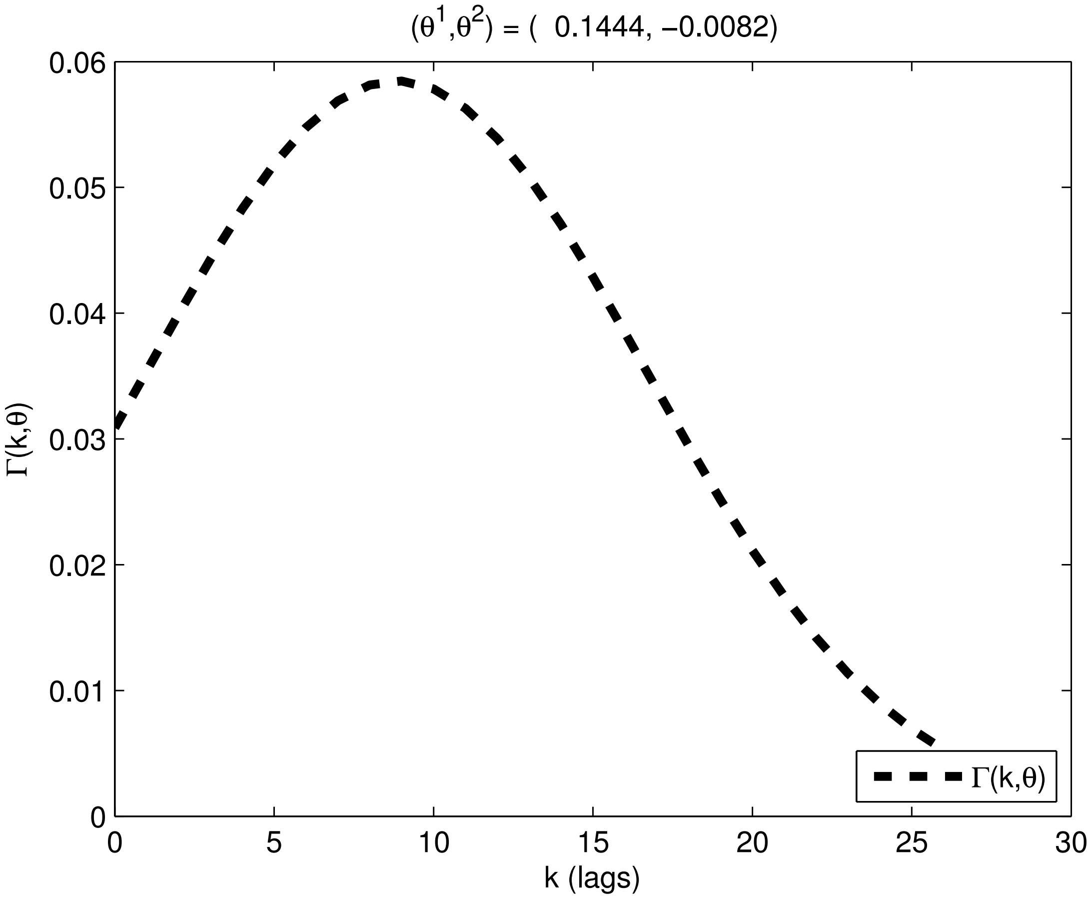


Figure 2: Typical Optimism Weights



(θ^1, θ^2) are the hyper-parameters of the weights function $\Gamma(k, \theta)$ defined in equation (2).