Journal of Forecasting, *J. Forecast.* **36**, 497–514 (2017)
Published online 21 November 2016 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/for.2450

Incorporating the Beige Book into a Quantitative Index of Economic Activity

NATHAN S. BALKE,[1,2*] MICHAEL FULMER[1] AND REN ZHANG[1]

[1] *Department of Economics, Southern Methodist University, Dallas, TX, USA*
[2] *Federal Reserve Bank of Dallas, TX, USA*

ABSTRACT
We apply customized text analytics to the written description of economic activity contained in the Beige Book (BB) in order to obtain a quantitative measure of current economic conditions. This quantitative BB measure is then incorporated into a dynamic factor index model that also contains other commonly used quantitative economic data. We find that at the time the BB is released it has information about current economic activity not contained in other quantitative data. This informational advantage is relatively short lived. By 3 weeks after their release date, 'old' BBs contain little additional information about economic activity not already contained in other quantitative data. Copyright © 2016 John Wiley & Sons, Ltd.

KEY WORDS Beige Book; text analysis; dynamic factor model

INTRODUCTION

The Beige Book (BB) is a written description of economic conditions in each of the 12 district banks of the Federal Reserve System. It is released eight times a year, roughly 2 weeks before the FOMC meeting. Its release is typically greeted with interest in the financial press. Is this attention warranted?[1] Does the BB contain information about current economic conditions? If so, can one use this information in a systematic manner? Can the information in the BB be combined with other quantitative information to provide a clearer picture of current economic conditions? In order to answer these questions, one must convert the qualitative information in the BB into quantitative information.

In this paper, we apply customized text analytics developed by Fulmer (2014) to extract quantitative information about current economic conditions from the BB. These quantitative BB indices are strongly correlated with current-quarter real GDP. We then incorporate a quantitative BB measure into a revised version of the Aruoba *et al.* (2009) (ADS) index model of daily economic activity. The ADS index is based on a dynamic common factor model that extracts a daily economic activity factor from six economic indicators: weekly jobless claims, monthly industrial production, personal income less transfers, employment, manufacturing trade growth and quarterly real GDP growth. We find that when using the full sample of final release data the incorporation of the BB index into the ADS model has little effect on the estimated daily index of economic activity. We find this somewhat comforting. Our priors are that, once the full sample of quantitative information is available, written information in the BB is unlikely to cause one to rewrite economic history. However, when conducting a 'real-time' analysis of the last three recessions, we find that the BB provides information about current economic activity not included in the other quantitative indicators present in the ADS index. Specifically, on the day that the BB is released, including the BB typically moves the ADS index towards its full-sample estimate ('the truth'). The informational contribution of the BB is typically larger during recessionary periods than in normal times. By about 3 weeks after the BB's release date, however, the ADS indices with or without the BB are very similar; the other information in the ADS index dominates whatever information the 'old' BB contains about current economic activity.

The paper is organized as follows. The next section contains a brief description of the BB and a brief review of the previous literature on quantifying the BB. The third section describes the customized text analytics applied to the BB. The fourth section describes how we incorporate the quantified BB into the dynamic factor model underlying the ADS index. The fifth section compares the properties of the ADS index with and without the BB. The sixth section concludes.

BACKGROUND

The BB is largely based on information that each of the 12 individual Federal Reserve District Banks gather from 'contacts' in their district. These 'contacts' are survey respondents, members of bank board of directors, news reports,

*Correspondence to: Nathan S. Balke, Department of Economics, Southern Methodist University, Dallas, TX, USA. E-mail: nbalke@smu.edu
[1] There is some skepticism about the usefulness of the BB. Alan Blinder (1997) somewhat disparagingly refers to the Fed's use of anecdotal information as 'ask your uncle' means of gathering information.

Copyright © 2016 John Wiley & Sons, Ltd

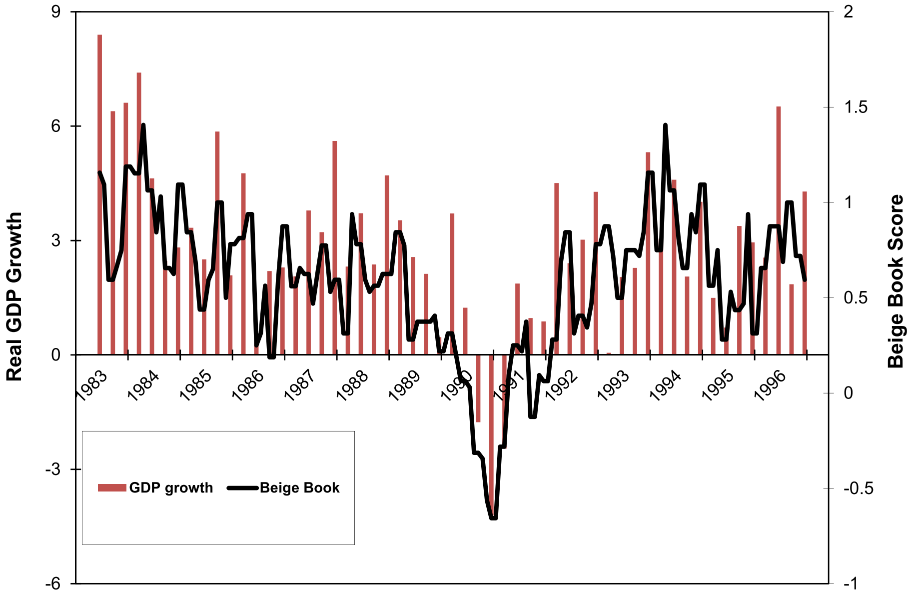

Figure 1. Real GDP growth and Balke–Petersen BB index (average of national sectors). [Colour figure can be viewed at wileyonlinelibrary.com]

and even officially released economic data. The information from these contacts is then summarized in the form of a written document. In addition to a written summary provided by each of the 12 districts, there is a national summary section in the BB. This is written by one of the district banks and is ostensibly based on the 12 district write-ups (the task of writing this summary is rotated among the 12 district banks). The information in the BB is compiled during the period up to the 'closing date', which is roughly 10 days prior to the official release date. For example, the BB that was released on 3 September 2014 had a closing date of 22 August 2014, and the BB that was released on 16 July 2014 had a closing date of 7 July 2014. The BB has been publicly released 2 weeks before the FOMC meeting since June 1983. Before that it was known as the 'Red Book' and was not publicly available. When we conduct our empirical analysis, we use both the Beige and Red Books.

Literature

Balke and Petersen (1998, 2002) were the first to attempt to quantify the information in the BB. They 'manually' read and scored each of the BBs from 1983 to 1997. Various features of the BBs were scored on a −2 to 2 scale by 0.5 increments.[2] The quantitative scores were then compared to actual real GDP growth. Figure 1 displays one of the Balke–Petersen BB indices and real GDP growth over their sample. Clearly, from Figure 1, the Balke–Petersen index tracks real GDP growth pretty well. In general, Balke and Petersen found that the BB had predictive content for current-quarter real GDP above and beyond other quantitative information available to analysts at the time the BB was released. On the other hand, Balke and Petersen found that the predictive content of the BB for next-quarter real GDP growth was substantially lower than for current-quarter real GDP growth. Follow-up studies that have used human readers/scorers (see Fettig *et al.*, 1999; Balke and Yucel, 2000; Ginther and Zavodny, 2001; Zavodny and Ginther, 2005) have also generally found that the BB contains information about current economic conditions both at the national and district level, but that the BB loses its informational advantage as other more recent data are released.

While the studies based on human readers of the BB suggest that there is information in it that could be useful for 'nowcasting' current economic activity, little effort has been given to updating these original studies of the quantified BB. There are several problems with employing human readers and scorers of the BB. First, the time it takes to read and score the BB makes it costly to update on a continual basis. Second, it is difficult to get a consistent reading of the BB over longer time periods—readers may vary over time and even the same reader's interpretation can change over time. These two problems make it difficult to replicate and systematically update the human reader BB indices.

Given the difficulties with 'manually' reading and scoring the Beige Book, there have been a few attempts to apply simple textual analysis to it. Payne (2001) developed a list of key word combinations and gave each a numerical score (−1 to 1 by 0.5 increments). He then counted the number of times these word combinations occurred in each district summary sentence. He found that the BB had strong predictive content for current quarter real GDP growth and for the Coincident and Leading Economic Indicators. Armesto *et al.* (2009) applied off-the-shelf text analytics software, *Diction*, to the BB. Using the word dictionary in *Diction*, Armesto *et al.* conducted a frequency count of words associated with 'optimism increasing' and 'optimism decreasing' to construct an optimism index ('optimism increasing') and a pessimism index ('optimism decreasing'). They used mixed frequency estimation (MIDAS) to account for

[2] They scored the National Summary, the four major sectors discussed in the National Summary, and the 12 district write-ups.

Copyright © 2016 John Wiley & Sons, Ltd

1099131x, 2017, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/for.2450 by Plymouth University, Wiley Online Library on [08/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the irregular spacing of the BB and found the it had predictive content for aggregate and district economic activity. Sadique *et al.* (2013) also used an off-the-shelf text analytics software, *GeneralInquirer*, to quantify the discussion of BB. Using the dictionary in *GeneralInquirer*, they extract the 'tone' of the BB, i.e. whether the discussion is positive/negative or increasing/decreasing. They found the 'tone' of the BB to be associated with phases of the business cycle as well as with changes in financial market volatility and volume.

TEXT MINING THE BEIGE BOOK

While the above studies go beyond simple word counts and are useful first attempts at quantifying the BB, our view is that they fall short of the full potential of applying text analytics. The ultimate goal of text analytics is to have the computer read the BB in a way similar to a sophisticated human reader. This would involve looking for word combinations or patterns that are of specific interest to the 'reader' and using application-specific word dictionaries. It would take advantage of grammatical relationships between words to extract meaning from the text. Finally, it would read the BB in 'context';i.e. it would use information outside the BB in order to interpret the discussion inside it.

In this application, we employ a customized text analysis of the BB. Our analysis builds on the work of Fulmer (2014). Fulmer develops a suite of algorithms to extract information from the BB about the state of current economic activity. These algorithms have many of the features that a sophisticated human reader uses when reading the BB: employing a word dictionary relevant for describing economic activity; searching for word combinations or patterns that are indicative of current economic activity; using grammatical structure to tease out economically meaningful relationships between words. The output of the text analysis will be five different BB indices. Three of the indices search for words or word patterns that have predictive content for real GDP growth. A fourth uses the grammatical structure of the text combined with a specialized word dictionary. The fifth is an ensemble index that includes elements of all four basic text analytic indices.

Initial filtering of the Beige Book documents

In this application, we use the discussion in the National Summary section of the BB. This is typically 4–5 pages long and represents a summary of the discussion in all the twelve Federal Reserve districts. When looking at the 393 Beige (and Red) Book documents produced from May 1970 to September 2014, there are over 13,000 unique character terms in the National Summary section. To extract meaningful information from the BB documents, we first filter the BBs by stripping each report of its punctuation and then extracting specific parts-of-speech from each sentence. The parts-of-speech included in the BB documents are limited to nouns, verbs, adjectives and adverbs. Furthermore, the stems of the words are used rather than the raw character strings in order to reduce the dimensionality of the terms and to combine morphologically similar character strings into concept groupings. Thus, character strings like 'Increasing', 'increasing', 'increased', 'increases', 'increase' are all collapsed to the term 'increas'. This filtering reduces the number of unique character terms to around 5000. Finally, we apply a custom BB dictionary. The BB dictionary consists of a list of accepted terms to be kept in the document; the rest of the terms are to be removed. Each of the close to 5000 unique nouns, verbs, adjectives and adverbs within the BB documents was individually accepted or rejected based on whether the term referred to or indicated the presence (or lack) of economic activity. After application of the dictionary filter, the number of terms in the BB documents is reduced to roughly 10% of its original size. The remaining terms represent the raw data to which we will apply statistical methods in order to extract information from the BB. These raw data we call the term document matrix, A, whose dimension is $n \times m$, where n is the number of terms and m is the number of documents (observations). The elements in the matrix are just counts of the number of times a particular term shows up in a particular document.

Five Beige Book indices

Despite the substantial reduction in the number of terms across documents achieved by filtering, the number of unique terms is still too large to provide clear signals about current economic activity. Furthermore, we want to extract the terms that are most relevant for the current growth rate in economic activity. We consider four alternative approaches to extract this information from the filtered document data.

Singular value decomposition (SVD) index

One way to reduce the dimension of terms to be considered is to apply the SVD to the term document matrix, A. The approach here is similar in spirit to principal component analysis. The SVD is given by $A = UDV'$, where $U_{(n\times n)}$ and $V_{(m\times m)}$ are orthonormal matrices, and $D_{(n\times m)}$ is a matrix with values along the diagonal and zeros off the diagonal. We take the columns of V associated with the k largest values of the D matrix. The dimensional choice, k, for the singular value projection of the term document matrix is a tuning parameter chosen by the researcher. Lower-dimension SVDs are better suited for out-of-sample prediction, while higher-dimension decompositions will yield better in-sample fits of the data. In this application, we chose a dimension of $k = 15$. Fulmer (2014) found

Copyright © 2016 John Wiley & Sons, Ltd

1099131x, 2017, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/for.2450 by Plymouth University, Wiley Online Library on [08/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table I. List of grammatical dependencies employed in the creation of BB atomic fact index

1	NSUBJ | NSUBJPASS	Subject noun
2.	NSUBJ | NSUBJPASS	Subject verb
3.	DOBJ	Direct object
4.	AMOD	Adjectival modifier
5.	(DOBJ) AMOD	Adjectival modifier (to the DOBJ modifier)
6.	ADVMOD	Adverbial modifier
7.	XCOMP	Open causal complement
8.	(XCOMP) DOBJ	Direct object (to the XCOMP modifier)
9.	(XCOMP) AMOD	Adjectival modifier (to the XCOMP DOBJ modifier)
10.	(XCOMP) ADVMOD	Adverbial modifier (to the XCOMP modifier)
11.	NEG	Negation modifier

that an SVD dimension between 10 and 20 balanced the trade-off between in-sample and out-of-sample performance reasonably well. These 15 'components' were then used in a regression of real GDP growth, the coefficients of which were used as weights on the components in the SVD index.

Positive/negative (PN) index
Here we construct an index similar to the 'optimism' and 'pessimism' indices created by *Diction* or the 'positive/negative' indices created by *GeneralInquirer*. Every term in a BB document was classified into one of three categories: positive, negative or neutral. For example, 'grow' would be classified as (+1), whereas 'decreas' would be classified as (−1) and 'unchang' would be classified as (0). The number of 'positive' terms and the number of 'negative' terms in a document were then summed to create indices of positivity and negativity for each document. These aggregate indices were then standardized by dividing by the total number of terms in the document. Having a positive and negative score for each BB, these were then used as independent variables in a real GDP regression. The coefficients from that regression formed the weights in the PN index.

Key word index
It is possible that the frequency of certain words might provide substantial information about economic activity.[3] Given the large body of words in the BB document, however, it may not be possible to identify these key words in advance. Thus the key word index searches for words whose frequency of usage appears to be related to real GDP growth. Given that there are close to 1300 raw terms even after the initial filtering, we first reduce these to the top 200 terms based on an entropy weighted frequency count (see Fulmer (2014)). This weighting assigns relatively greater weight to words that have high counts in a few documents and low counts in the remaining documents. Words whose frequency across the documents is relatively uniform provide little information about changes in economic activity and hence will receive low weight. Of these 200 terms, we then search for the set of terms that is most closely related to real GDP growth.[4] It turns out that the best subset consists of the frequency count of the following 10 terms: 'increas', 'strong', 'declin', 'cost', 'weaken', 'eas', 'vacanc', 'brisk', 'strike' and 'borrow'.

Atomic fact extraction (AFE) index
Unlike the previous three indices that essentially look for patterns in word counts across the BB documents, the AFE uses grammatical relationships between words to extract information from the text. The simplest grammatical relationship would be the subject noun and the verb from each sentence. However, other grammatical relationships in a sentence might prove useful as well. Table I lists the set of 11 grammatical dependencies considered. Similar to the positive/negative index discussed above, every noun, verb, adverb and adjective in the BB dictionary was assigned a positive, negative or neutral value (+1, −1 or NA, respectively). Using these values, the atomic fact score of the sentence clause is created by multiplying the classifiers together. For example, a sentence with the relationship 'output rises' would score $(+1) \times (+1) = 1$, while another sentence with 'layoffs increas' would score $(-1) \times (+1) = -1$. After obtaining the score for each individual sentence clause, the atomic fact index of the document is obtained by summing the scores of the individual sentence clauses and then dividing by the maximal score possible for the document. This standardization ensures that longer documents do not receive larger index values simply because they contain more sentence clauses than shorter documents.

Ensemble index
Here we combine the elements in the previous four indices. Specifically, we take the components identified in the text analysis—the SVD components, the positive and negative indices, the 10 'key' words and the AFE index—and put all

[3] For example, the *Economist*'s recession index is a count of the number of times 'recession' appears in a select group of news publications.
[4] We use the subset regression algorithm in R to determine the best model (see Miller, 2002; Gatu, 2006). This algorithm searches for the 'best' model for a given number of terms (here from 5 to 20 terms). The optimal number of terms is determined by minimizing the Bayesian information criterion.

Copyright © 2016 John Wiley & Sons, Ltd

1099131x, 2017, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/for.2450 by Plymouth University, Wiley Online Library on [08/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

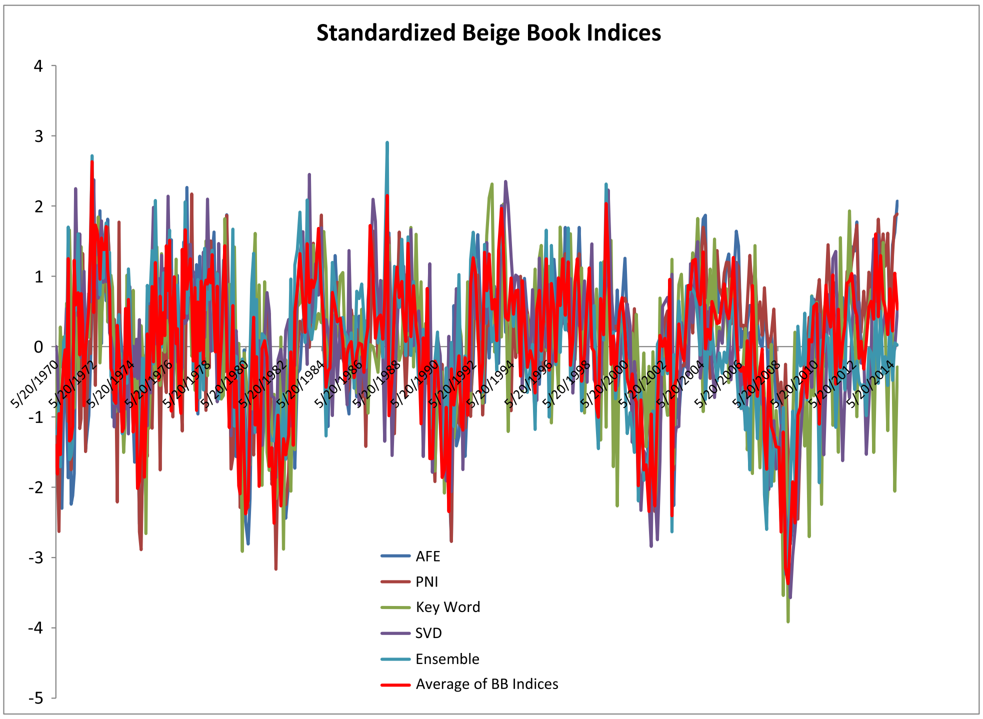

Figure 2. Standardized Beige Book Indices (dates in figure are in m/d/yyyy format). [Colour figure can be viewed at wileyonlinelibrary.com]

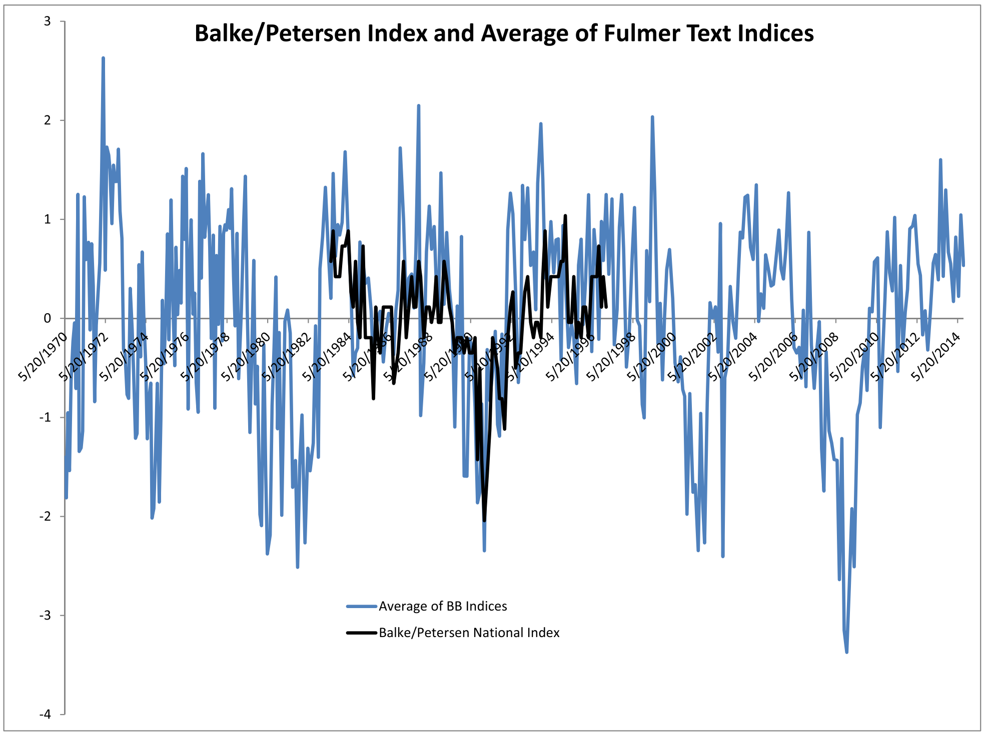

Figure 3. Balke–Petersen index and average of Fulmer text indices (dates in figure are in m/d/yyyy format). [Colour figure can be viewed at wileyonlinelibrary.com]

of these in a single regression with real GDP growth as the dependent variable. The coefficients from the regression determine the weights that the various components receive in the ensemble index.

Quantitative properties of the Beige Book documents

Figure 2 displays the five BB indices. To make them comparable, we standardize each index by subtracting their sample means and dividing by their sample standard deviations. Figure 2 also plots the simple average of the five BB series (similarly standardized). From the figure, it is clear that these indices move together and share many of the same cyclical qualities. The contemporaneous correlations among the five series range from a low of 0.34 (between the PN index and the key word index) to a high of 0.87 (between the AFE and the PN index). To see how the text analytics

1099131x, 2017, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/for.2450 by Plymouth University, Wiley Online Library on [08/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table II. Real GDP regression on BB components

Dependent variable: Real GDP growth					
Independent variable	BB index				
	AFE	PN	Key word	SVD	Ensemble
Constant	0.440****	1.141***	0.743****	−0.284	0.159
AFE index	2.504****				−0.133
Positive index		2.058			−0.975
Negative index		−9.418***			0.529
‘increas’			2.415***		0.239
‘strong’			3.450****		−1.801
‘declin’			−3.988****		−2.306
‘strength’			1.360****		1.581****
‘weaken’			−1.233****		−0.863***
‘eas’			−0.994****		−0.920****
‘recover’			0.680**		0.850***
‘strike’			−0.655***		−0.716***
‘borrow’			−0.743**		−0.602***
‘recess’			−0.770***		−0.692**
svd1				−0.284	−0.404*
svd2				−0.327****	−0.276***
svd3				0.289***	0.341***
svd4				0.151*	0.025
svd5				−0.406****	−0.204
svd6				−0.304****	−0.093
svd7				0.104	0.096
svd8				−0.244**	−0.062
svd9				−0.034	−0.196*
svd10				−0.097	−0.047
svd11				−0.207*	0.012
svd12				0.038	−0.103
svd13				−0.106	0.063
svd14				0.173*	0.249*
svd15				−0.043	−0.096
Adj. R^2	0.134	0.122	0.349	0.265	0.361
p-value for excluding:					
Key words			2.2E−16****		5.9E−10****
SVDs				2.2E−16****	0.094*

Note: Asterisks indicate the following levels of significance: ****0.001; ***0.01; **0.05; *0.1.

Beige Book score compares with scores based on human readers, Figure 3 displays the average of the five BB indices with the standardized Balke–Petersen National Summary index. Both indices share the same cyclical pattern and have a correlation coefficient of 0.57. This suggests that both computer reading and human reading of the BB yield similar interpretations.

To get a sense of how well the Beige Book indices capture changes in aggregate economic activity, Table II contains the results of regressing real GDP growth against the underlying components making up the BB indices. For all five indices, the BB components are jointly statistically significant. Not surprisingly, given the substantially larger number of parameters, the SVD, key word and ensemble indices have better in-sample fit than the AFE and PN indices. Including lags of real GDP growth in the regression does not eliminate the joint significance of the BB components in the real GDP regression (see Table III).

DAILY INDEX OF ECONOMIC ACTIVITY CONTAINING THE BEIGE BOOK

While the quantitative BB indices appear to reflect current economic activity, most analysts are not likely to use the BB in isolation. Here we add the average of our five BB indices to the revised version of the Aruoba *et al.* (2009) index of daily economic activity. The most recent version of this model is maintained by the Federal Reserve Bank of Philadelphia. One challenge in combining the BB with other economic indicators is the irregular frequency of the BB's release. The ADS model is particularly attractive for our application as it can readily accommodate data of different frequencies and, in particular, the irregular frequency of the BB.

ADS common factor model

The ADS model is a common factor model in which the underlying state variable reflects daily economic activity and the common comovement of the observables is the result of time aggregation of this underlying daily state variable. Indicators include observations of weekly initial claims for unemployment insurance (JC), quarterly real

Copyright © 2016 John Wiley & Sons, Ltd

1099131x, 2017, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/for.2450 by Plymouth University, Wiley Online Library on [08/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table III. Real GDP regression on BB components

Dependent variable: Real GDP growth					
	BB index				
Independent variable	AFE	PN	Key word	SVD	Ensemble
Constant	0.169****	0.285***	0.233*	−0.608	−0.362
AFE	0.683****				−0.393
Positive		0.673			−0.815
Negative		−2.046			1.818
‘increas’			1.543***		1.875
‘strong’			2.162***		−2.577
‘declin’			−1.859****		−1.610*
‘strength’			0.602*		1.091***
‘weaken’			−0.630***		−0.452*
‘eas’			−0.609***		−0.626***
‘recover’			0.596***		0.788****
‘strike’			−0.303**		−0.369**
‘borrow’			−0.401**		−0.323
‘recess’			−0.253		−0.231
svd1				−0.270	−0.317*
svd2				−0.157***	−0.178**
svd3				0.095	0.247**
svd4				0.105*	0.034
svd5				−0.201****	−0.029
svd6				−0.156**	−0.120
svd7				0.065	0.143
svd8				−0.154**	−0.059
svd9				−0.058	−0.233***
svd10				−0.099	−0.085
svd11				−0.007	0.135
svd12				0.042	−0.110
svd13				−0.066	0.182*
svd14				0.125	0.235**
svd15				−0.056	−0.065
Adj. R^2	0.157	0.513	0.571	0.547	0.583
p-value for excluding:					
4 lags of GDP	2.2E−16****	2.2E−16****	2.2E−16****	2.2E−16****	2.2E−16****
Key words			3.7E−09****		1.7E−05****
SVDs				9.5E−0.5****	0.024**

Note: Asterisks indicate the following levels of significance: ****0.001; ***0.01; **0.05; *0.1.

GDP (GDP), monthly industrial production growth (IP), growth in real personal income less transfer receipts (I), monthly real manufacturing and trade sales (MT), and monthly employees and non-agricultural payrolls (EM). To these six indicators, we add our BB index. Following the Philadelphia Fed version of the ADS index, we log difference all the variables except initial jobless claims and the BB index. Accordingly, the vector of observables is expressed as

$$y_t = (\mathrm{JC}_t,\ \Delta\mathrm{GDP}_t,\ \mathrm{BB}_t,\ \Delta\mathrm{IP}_t,\ \Delta I_t,\ \Delta\mathrm{MT}_t,\ \Delta\mathrm{EM}_t)' \tag{1}$$

In our estimation, we will standardize all the observable indicators by subtracting their sample means and dividing by their sample standard deviations.[5]

Let x_t denote underlying economic activity on day t. x_t is assumed to follow an AR(p) process:

$$x_t = \rho_1 x_{t-1} + \rho_2 x_{t-2} + \ldots + \rho_p x_{t-p} + e_t \tag{2}$$

where e_t is a white noise innovation with variance so that x_t has a unit unconditional variance. If all the data were observed daily, the indicator i, y_t^i, depends on x_t and lags of y_t^i:

$$y_t^i = c_i + \beta_i x_t + \gamma_{i1} y_{t-1}^i + \ldots + \gamma_{in} y_{t-n}^i + u_t^i \tag{3}$$

where u_t^i are contemporaneously and serially uncorrelated innovations. The parameter β_i is the factor loading of daily economic activity on the observable variable.

In practice, the indicators are not observed at a daily frequency. As the indicators are flow variables and are observed at a lower frequency than daily, one must temporally aggregate the effect of the daily economic variable

[5] When conducting our real-time analysis below, we will standardize by the real-time sample means and standard deviations.

Copyright © 2016 John Wiley & Sons, Ltd

1099131x, 2017, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/for.2450 by Plymouth University, Wiley Online Library on [08/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

on the observables. As a result, the observations of indicator i, $\tilde{y}_t^i$, are intra-period sums of the corresponding daily values:

$$\tilde{y}_t^i = \begin{cases} \sum_{j=0}^{D_i-1} y_{t-j}^i & \text{if } \tilde{y}_t^i \text{ is observed} \\ \text{NA} & \text{otherwise} \end{cases} \tag{4}$$

and D_i denotes the number of days since the last observation period. For all but most the recent time periods, we assume $\tilde{y}_t^i$ is observed on the last day of the relevant period. For recent periods in which official data for the relevant time period have yet to be released, we treat $\tilde{y}_t^i$ as missing (or NA).

Instead of combining equations (2) and (3) directly, cumulator variables, as proposed by Harvey (1989), are used to handle the temporal aggregation of the flow variables. By cumulating values of the underlying business condition x_t, the cumulator variables summarize all the information needed to construct aggregated flow variables. We define the cumulators C_t^f, $f \in \{\text{W, M, Q, BB}\}$ as

$$C_t^f = \zeta_t^f C_{t-1}^f + x_t \tag{5}$$

$$= \zeta_t^f C_{t-1}^f + \rho_1 x_{t-1} + \ldots + \rho_p x_{t-p} + \epsilon_t \tag{6}$$

where C_t^{W}, C_t^{M}, C_t^{Q}, C_t^{BB} denote the weekly cumulator, monthly cumulator, quarterly cumulator, and the cumulator for BB index, respectively. Additionally, ζ_t^f is an indicator variable, defined as

$$\zeta_t^f = \begin{cases} 0 & \text{if } t \text{ is the first day of a period} \\ 1 & \text{otherwise} \end{cases} \tag{7}$$

For weekly, monthly and quarterly variables, the first day of the period corresponds to the first day of the week, month and quarter, respectively. For the BB variable, we take the first day of the period to be the day after the closing date of the previous BB. Hence the BB will reflect the accumulated daily economic activity from the day after the closing date of the previous BB to the closing date of the current BB.

Based on the definition of the cumulators, the measurement equation for a generic flow variable $\tilde{y}_t^i$ can be written as

$$\tilde{y}_t^i = \begin{cases} c_i^* + \beta_i C_t^i + \gamma_{i1}\tilde{y}_{t-D_i}^i + \ldots + \gamma_{in}\tilde{y}_{t-nD_i}^i + u_t^{*i} & \text{if } \tilde{y}_t^i \text{ is observed} \\ \text{NA} & \text{otherwise} \end{cases} \tag{8}$$

where u_t^{*i} is an idiosyncratic shock to indicator i. The above model can be put into state-space form (see Appendix for details) and estimated by maximum likelihood. For a given parameter vector, the likelihood can be evaluated using a Kalman filter and an estimate of the unobserved daily index can be obtained using the Kalman smoother. Note that factor loading for the daily economic activity state variable in the BB observation equation (β_{BB}) is positive and statistically significant, implying that observations on the BB variable can shed some light on the (directly) unobserved daily factor (see Table A1 in the Appendix).

PERFORMANCE OF BEIGE BOOK ADS INDEX

Full-sample ADS indices

Figure 4 presents the full-sample (from May 1970 to September 2014) estimates of the ADS index for the model with and without the BB.[6] For reference, real GDP growth is also displayed. As one can see, over the full sample both ADS indices track real GDP growth very well. Furthermore, the difference between the full-sample ADS index with and without the BB is relatively small; both indices are telling the same story about the underlying daily economic activity. To get a closer look at the relative behavior of the ADS indices, Figure 5 just displays the period since 2007. Again, both the ADS with the BB and without the BB track real GDP growth relatively well. There are some small differences in the two indices perhaps worth noting. In early 2008, the ADS with the BB suggests a slightly larger decline in economic activity than the ADS without the BB, suggesting that economic activity was slightly weaker than implied by the other quantitative indicators. In September 2008, due to sharp one-time declines in industrial production and manufacturing and trade data, the ADS indices display a sharp drop in September and then a sharp but temporary ‘rebound’ in October 2008. However, for the ADS index with the BB this ‘bounce back’ is slightly less pronounced and presages the continued deterioration in the ADS indices later in 2008 and early 2009. These

[6] These are based on the Kalman smoothed estimate of the state variables using the full sample to estimate the parameters in the model as well as in the Kalman smoother.

Copyright © 2016 John Wiley & Sons, Ltd

1099131x, 2017, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/for.2450 by Plymouth University, Wiley Online Library on [08/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

small differences notwithstanding, both indices are telling very similar stories about the underlying daily economic activity factor.

Figures 4 and 5 suggest that when one has full-sample information about all the other indicators, the BB does not provide much additional information about current economic activity. We find this result unsurprising. We would not really expect the BB to dramatically change our view of economic activity over time once the final releases of real GDP, industrial production, employment growth and the other indicators were available. That the full-sample ADS indices with and without the BB are similar suggests that we are not 'reading too much into BB'.

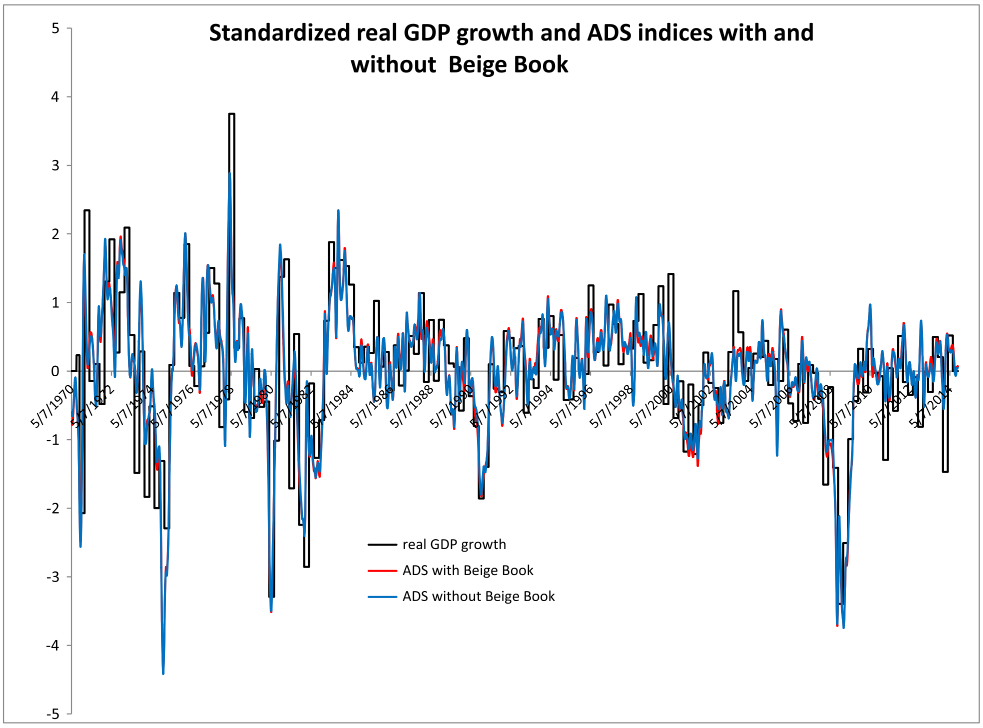

Figure 4. Standardized real GDP growth and ADS indices with and without BB (dates in figure are in m/d/yyyy format). [Colour figure can be viewed at wileyonlinelibrary.com]

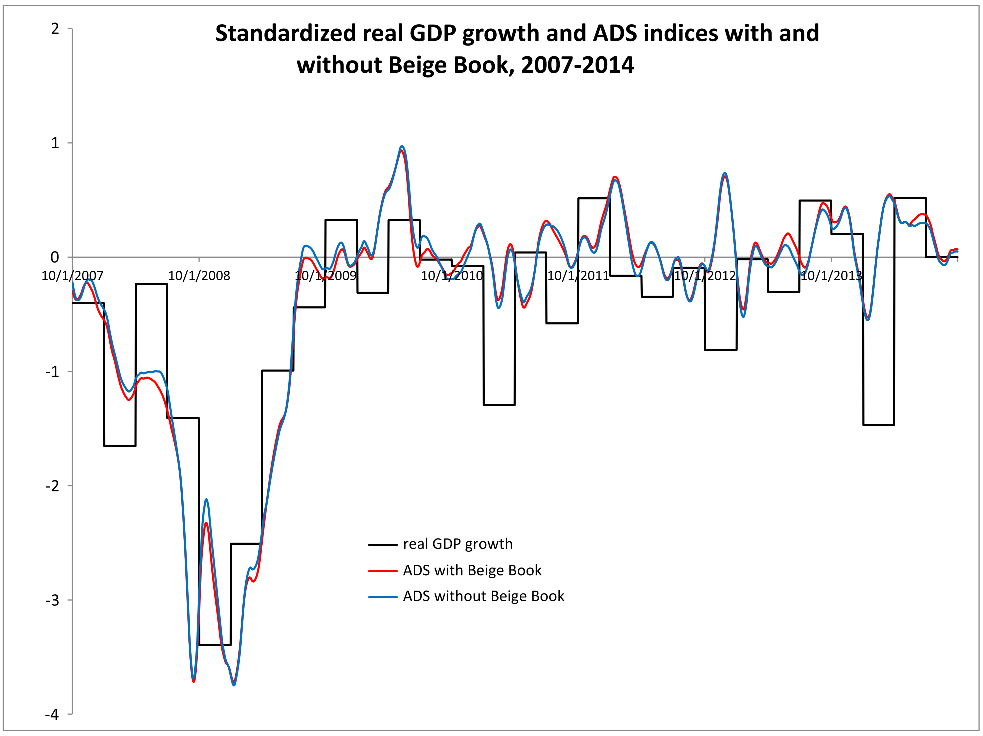

Figure 5. Standardized real GDP growth and ADS indices with and without BB, 2007-2014 (dates in figure are in m/d/yyyy format). [Colour figure can be viewed at wileyonlinelibrary.com]

Copyright © 2016 John Wiley & Sons, Ltd

1099131x, 2017, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/for.2450 by Plymouth University, Wiley Online Library on [08/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table IV. Mean squared prediction errors for real-time ADS Indices with and without Beige Book information.

	Mean squared prediction error				
Sample	ADS model with Beige Book (1)	ADS model without Beige Book (2)	Ratio of (1) and (2) (3)	C-W t-stat for forecast difference (4)	Number of forecasts in sample (5)
	Panel A. Forecast comparison on dates that Beige Book is released				
3/1/1990-12/31/1991	0.2710	0.3396	0.80	1.9176*	15
12/1/2000-3/31/2002	0.2137	0.2455	0.87	1.3139	11
10/1/2007-12/31/2009	1.5286	1.836	0.83	2.4769*	18
	Panel B. Forecast comparison using entire sample				
3/1/1990-12/31/1991	0.2638	0.287	0.92	4.994*	671
12/1/2000-3/31/2002	0.2567	0.2592	0.99	0.869	486
10/1/2007-12/31/2009	1.2304	1.3101	0.94	5.3106*	823

Note: The variable being forecasted is the full sample estimate of the Index of Economic Activity implied by the ADS model with Beige Book.

Real-time examination of the ADS indices

While the BB might not cause analysts to revise their view of economic history once all the data are available, where the BB might be useful is when only incomplete or noisy information is available. The standard quantitative data are typically available with a lag and are often subsequently revised. For example, preliminary estimates of real GDP are released roughly 1 month after the quarter has ended. Furthermore, the real GDP data are going to be revised twice over the subsequent 2 months before the ‘third’ estimate is released—nearly 3 months after the quarter has ended.[7] The BB with its relatively quick release date (roughly 10 days after the closing period) and text that is not subsequently revised may provide timelier and less noisy information about current economic conditions than some of the other economic indicators.

To investigate whether the BB is useful when there is only partial information, we conduct a ‘real-time’ examination of the ADS index with and without the BB.[8] In particular, we examine the last three recessions to determine whether including the BB can provide timely information about economic activity during these periods. To do this, we only use information that would have been available to analysts at the time both to estimate the ADS index models and then to update the estimate of economic activity. Specifically, we use data that were available up to a specified date before the recession to estimate the ADS model with and without the BB; this also includes estimating the BB indices using data that were available to analysts up to that specified time period.[9] Then, during the recession period, we update the inference about the underlying state variables using data as they become available in ‘real time’.[10] To obtain a quantitative assessment of the effect of including the BB information in the index of economic activity we calculate the real-time mean squared predictive error (MSPE), where the ‘target’ is the full-sample estimate of the ADS index. Here we compare both real-time ADS indices (with and without the BB) to the full-sample ADS index with the BB information.[11] In terms of state-space model described above, this amounts to calculating the difference between the Kalman smoothed value of the index based on the model estimated over the full sample and the Kalman filtered estimate of the model estimated with information just prior to the recession. To test whether the two real-time indices yield statistically different forecasts, we use the Clark and West (2007) MPSE-adjusted statistic to test whether the alternative forecasts are significantly different from one another.[12] Panel A of Table IV displays the MSPEs for

[7] Furthermore, periodically, the historical GDP data are revised, meaning that the ‘final’ estimate of GDP might be released years after the fact.
[8] For manufacturing and the personal income series, we use the real-time dataset compiled by Giusto and Piger (2014) which, in turn, was based on Chauvet and Piger (2008). The industrial production, real GDP, weekly claims for unemployment insurance and monthly employees and non-agricultural payrolls are from the Philadelphia Fed real-time website as well as from FRED (St Louis Fed).
[9] The estimation periods are: for the 2007–2009 recession data up to 30 September 2007; for the 2001 recession data up to 30 November 2000; and for the 1990–1991 recession data up to 28 February 1990.
[10] We do not re-estimate the entire factor model, but use the updated data and the Kalman filter/smoother to estimate the underlying daily economic activity factor. While it would be possible to re-estimate the model during the recession periods, in practice the original ADS model is only re-estimated annually. Our procedure to use the model as initially estimated at a pre-specified period before the recession approximates this practice.
[11] As a check, we also compared the real-time ADS indices to the full-sample ADS index, which does not include the BB. As the full-sample ADS indices either with or without the BB are so similar, it makes little difference when comparing the real-time ADS indices which full-sample ADS index we use.
[12] The null hypothesis is equal MSPE across the two models, while the alternative is that the ADS model with BB has a smaller MSPE. As suggested by Clark and West, we use standard-normal critical values as approximate critical values.

Copyright © 2016 John Wiley & Sons, Ltd

1099131x, 2017, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/for.2450 by Plymouth University, Wiley Online Library on [08/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

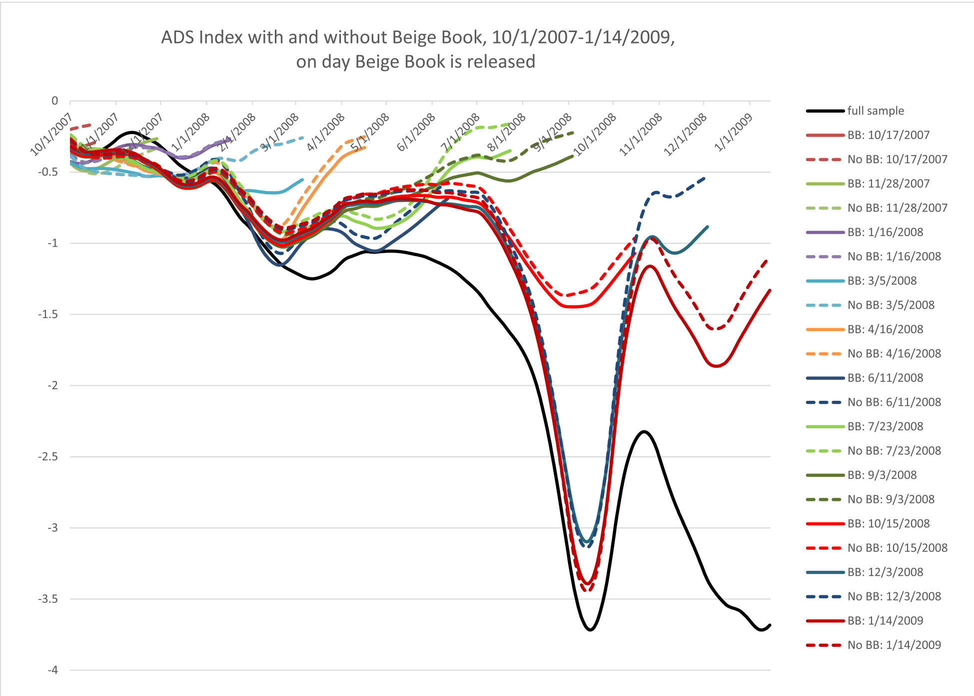

Figure 6. ADS Index with and without BB, 1 October 2007 to 14 January 2009, on day Beige Book is released (dates in figure are in m/d/yyyy format). [Colour figure can be viewed at wileyonlinelibrary.com]

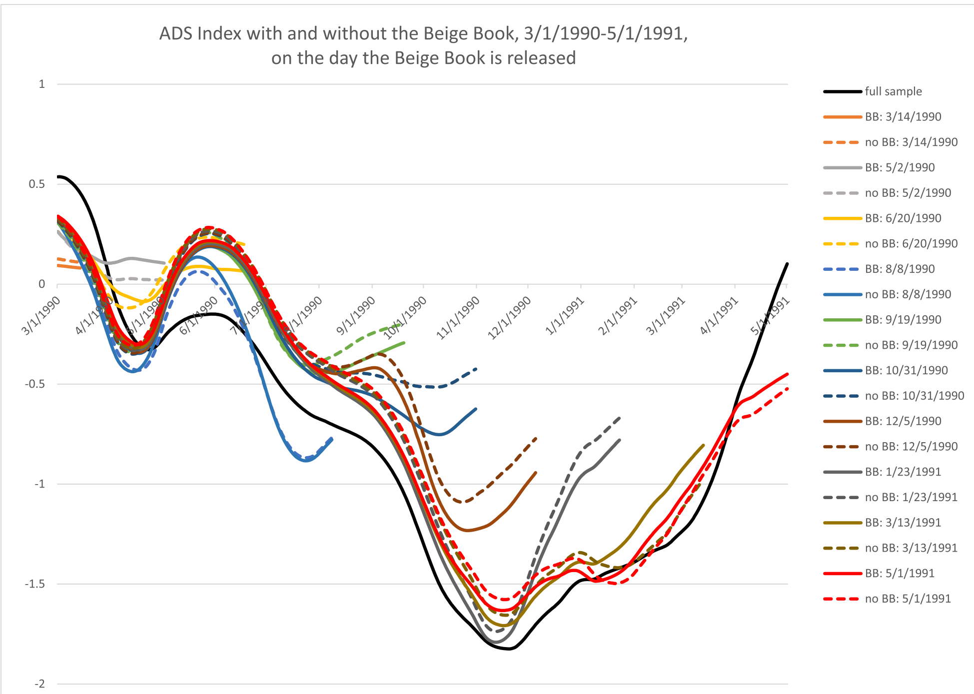

Figure 7. ADS index with and without the BB, 1 March 1990 to 1 May 1991, on the day the Beige Book is released (dates in figure are in m/d/yyyy format). [Colour figure can be viewed at wileyonlinelibrary.com]

the models with and without the BB for just the dates on which the BBs were released. Panel B of Table IV displays the MSPEs for the competing models where every daily observation is forecast over the respective sample periods.

From panel A, on the dates the BB is released, including the BB information generally lowers the squared prediction errors by 15% in both the 1990–1991 and the 2007–2009 recessions. These reductions are statistically significant. On the other hand, for the 2001–2002 period, including the BB did not result in a statistically significant lower MSPE. These results suggest that, at least for the 1990–1991 and 2007–2009 recessions, at the time the BB was released it could have provided additional information about current economic activity not contained in the other currently available data used in the ADS index. From panel B, one observes that the advantage of including the BB information

Copyright © 2016 John Wiley & Sons, Ltd

1099131x, 2017, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/for.2450 by Plymouth University, Wiley Online Library on [08/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

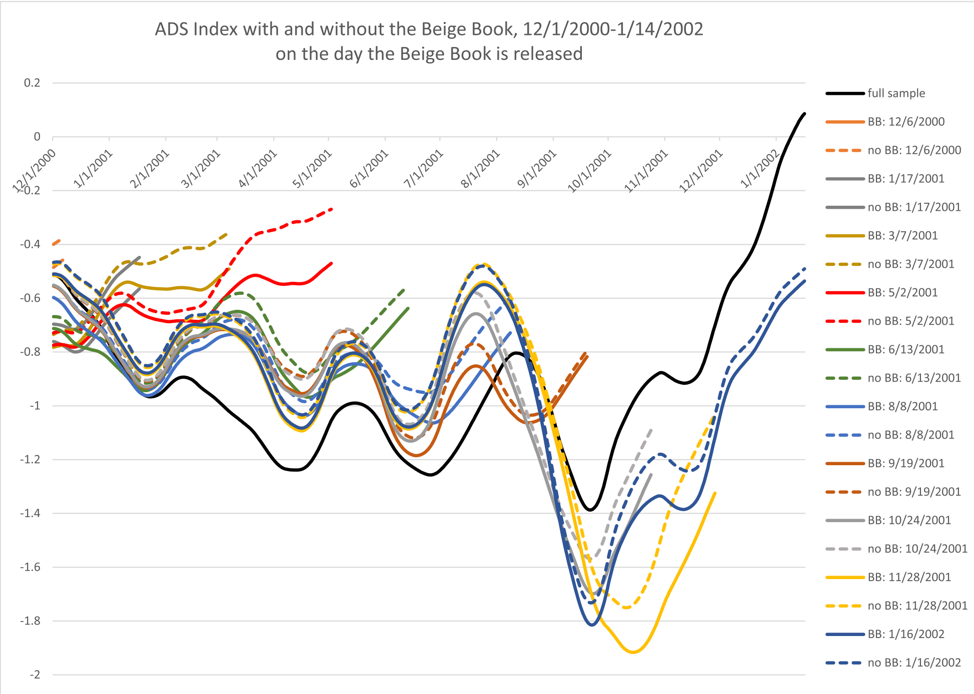

Figure 8. ADS index with and without the BB, 1 December 2000 to 14 January 2002 on the day the Beige Book is released (dates in figure are in m/d/yyyy format). [Colour figure can be viewed at wileyonlinelibrary.com]

appears to extend to non–BB release dates. When using every daily observation to compare forecasts, including the BB results in significantly improved forecasts (in terms of MSPEs) for the 1990–1991 and 2007–2009 recessions.

Figure 6 displays the various paths of the estimated ADS index with and without the BB based on information available on the dates that the BB was released. To keep Figure 6 from being too cluttered, we plot just the period from 1 October 2007 to 14 January 2009. The solid lines are the estimated ADS index for the model with the BB, while the dashed lines are for the model without the BB. For comparison, the black line in the figure is the estimated index based on the entire sample, which we take as the ‘true’ level of economic activity. Figure 6 suggests that on the BB release dates incorporating the BB tends to move the estimated ADS index towards its full-sample estimate. For example, take the ADS index based on information available on BB release dates: 5 March 2008, 6 April 2008, 11 June 2008 and 3 September 2008. Including the BB information moves the estimated daily index towards to the ‘truth’ by between 20 and 50 basis points. Note that these were periods well before the depth of the actual economic decline was evident, yet the BB is suggesting that economic activity was not as strong as suggested by the other available quantitative data. Similarly, including the BB information released on 3 December 2008 moves the index by nearly 50 basis points down towards the eventual full-sample estimate.

Figures 7 and 8 do a similar exercise as Figure 6 for the 1990–1991 and 2001–2002 recessions, respectively. In many of the periods early in these recessions, at the time the BB is released, using the BB information yields estimates of the ADS index closer to the eventual full-sample estimate. In the 1990–1991 recession, the BBs released on 20 June 1990, 19 September 1990, 31 October 1990 and 5 December 1990 move the ADS index towards its full-sample estimate. Note that a BB release does not always lead to more accurate revision of the underlying state of economic activity. For example, the BB released on 8 August 1990 has virtually no effect on the estimated index and the BB released on 2 May 1990 actually moves the index further away from the ‘truth’. Nonetheless, at the time that the BBs are released, many of the BBs early in this recession improve the estimate of current economic activity. For the 2001–2002 recession, while Table IV indicated that the BB did not improve forecasts of the full-sample ADS index, Figure 8 suggests that adding the BB information actually improves the forecast early in that recession. For example, the BBs released on 17 January 2001, 7 March 2001 and 2 May 2001 move the ADS index towards its full-sample estimate. On the other hand, in the period after the trough of the 2001 recession, including the BB tends to move the ADS index away from its full-sample estimate (see ADS estimates on 24 October 2001 and 28 November 2001), here overestimating the depth of that recession in real time.

Taken together, the results of Table IV and Figures 6–8 indicate that information in the BB could prove useful in real-time forecasting of the full-sample ADS index, particularly early in a recession. This also implies that including the BB might be useful in helping identify recession periods in real time. Berge and Jorda (2011) argue that periods where the (full-sample) ADS index falls below a threshold of −0.8 are useful for detecting recessions. While Figures 6–8 demonstrate that the −0.8 threshold applicable to the full-sample ADS index may not be applicable to

Copyright © 2016 John Wiley & Sons, Ltd

1099131x, 2017, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/for.2450 by Plymouth University, Wiley Online Library on [08/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

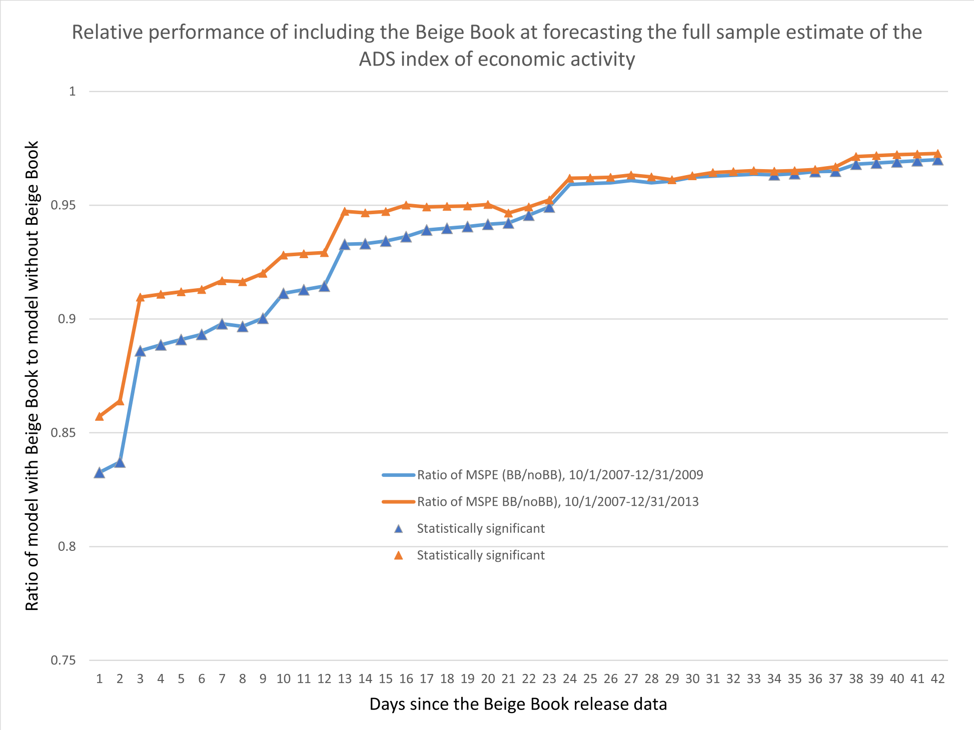

Figure 9. Relative performance of including the BB at forecasting the full sample estimate of the ADS index of economic activity. [Colour figure can be viewed at wileyonlinelibrary.com]

real-time detection of recessions, including the BB information does move the ADS index closer to the Berge and Jorda threshold sooner.[13]

How long does the BB's informational advantage last? As other quantitative data are released, the information contained in the BB about economic activity might be overshadowed by more recently released data. We can get a sense for how long the BB advantage persists by examining how the inclusion of the BB information affects the 'nowcasting' performance on the days after the BB release. As the number of days since the BB release increases, the more likely new quantitative data has arrived. By the day before the next BB is to be released (roughly 41 days since initial BB release), the BB information is nearly 6 weeks old and there have been at least five new releases of jobless claims, one new release of industrial production, employment, personal income less transfers and manufacturing/trade data, and a revision of the previous quarter GDP estimate, if not a release in the preliminary estimate of GDP.

Figure 9 displays the ratio of the two models' MSPEs (for predictions of the full-sample ADS index) from 0 to 41 days after the BB release. A ratio less than one implies the BB model has an informational advantage over the model without the BB. We use the Clark and West (2007) test to examine whether the forecasts are statistically different.[14] For the recession period (1 October 2007 to 31 December 2009), the BB's informational advantage is largest just after it is released (roughly a 17% improvement) and gradually fades as the number of days since the release date increases. By about 22 days, the difference is only 5% and no longer statistically significant. For the entire 1 October 2007 to 31 December 2013 sample, the improvement in forecasts of incorporating the BB is smaller in magnitude than during just the recession period (13% vs. 17% improvement on the BB release date) and by about 22 days the improvement is only 5%.[15] Thus, on average, by the time that the BB is 3 weeks old, it provides only a little additional information about the ADS index of economic activity not already included in the other quantitative indicators.

Real-time forecasting of current-quarter real GDP growth

In the previous section, we conducted a real-time analysis of the contribution of the BB to predicting the full-sample (final) estimate of the underlying daily index of economic activity. In this section, we examine the contribution of including the BB as an observation variable in the ADS model (along with the other quantitative variables) towards predicting the current quarter real GDP growth (final estimates).

[13] A more detailed examination of the ability of the ADS indices with and without the BB information to detect turning points in real time is left to future research. We thank an anonymous referee for suggesting this line of inquiry.

[14] Again, we examine the real-time forecast relative to the full-sample ADS index that includes the BB. The results are essentially unchanged if we compare the real-time forecasts to the full-sample ADS index without the BB.

[15] The improvement is, however, statistically significant throughout. The fact that the larger sample yields smaller improvements in magnitude yet is statistically significant is probably due to the larger sample size (50 vs. 18 release dates).

Copyright © 2016 John Wiley & Sons, Ltd

1099131x, 2017, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/for.2450 by Plymouth University, Wiley Online Library on [08/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

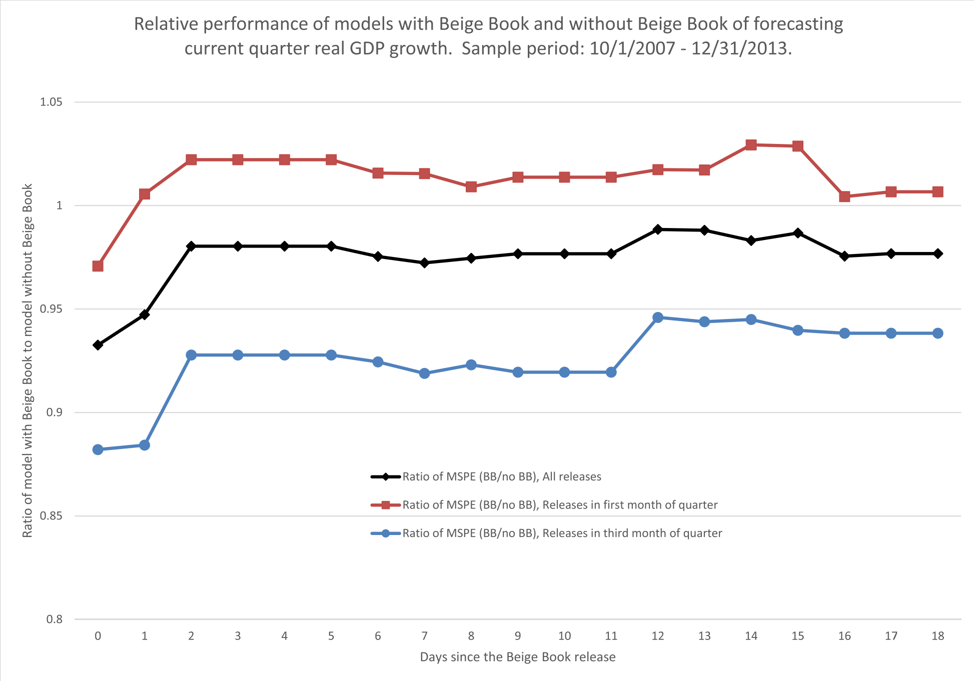

Figure 10. Relative performance of models with and without BB of forecasting current-quarter real GDP growth. Sample period: 1 October 2007 to 31 December 2013. [Colour figure can be viewed at wileyonlinelibrary.com]

As our focus is on the usefulness of the BB information in a quantitative index of economic activity and not on forecasting real GDP growth per se, we do not consider the large number of other forecasting models of real GDP growth. Instead, we use the ADS index model to forecast current quarter real GDP growth given the available information at a particular time. Specifically, we can use the observation equation for real GDP in the ADS model to forecast current-quarter real GDP, or

$$\tilde{y}^{GDP}_{t_Q} = c^*_{GDP} + \beta_{GDP} C^Q_{t_Q|t} + \gamma_{GDP,1} \tilde{y}^{GDP}_{t_Q - D_{GDP,1}} \tag{9}$$

where t is the time period in which the forecast is being made and t_Q is the last day of the current quarter.footnoteWe actually 're-standardize' real GDP growth by multiplying the model's prediction for standardized real GDP growth by the sample standard deviation of real GDP growth and then add back in the sample mean. We consider time periods in the current quarter so that $t < t_Q$. From equation (9), this largely amounts to forecasting the value of the quarterly cumulator of the underlying daily state variable at the end of the quarter, $C^Q_{t_Q|t}$.[16]

We calculate the ratio of the MSPEs for forecasts of the final estimates of current-quarter real GDP implied by the ADS index with and without the BB. Here we examine the period 1 October 2007 until 31 December 2013. We also break up the MSPEs for forecasts when the date of the forecast is in the first month of the quarter and when the date is in the third month of the quarter. The idea is that forecasts of current-quarter real GDP growth are likely to be worse when in the first month of the quarter than when in the third month of the quarter. In particular, BBs released in the first month of the quarter will contain information largely about previous-quarter real GDP, while BBs released in the third month of the quarter will contain information about the current quarter only.[17]

From Figure 10, we observe that, aggregating over all BB release dates, the ADS model with the BB improves forecasts of current-quarter real GDP growth relative to the model without the BB by around seven percentage points at the time that the BB is released. This informational advantage persists, although the magnitude of this advantage is relatively small.[18] Not surprisingly for BBs released in the first month of the quarter, the informational advantage is relatively small (and even negative). However, BBs released in the third month of the quarter improve MSPEs for current-quarter GDP growth by over 12%. Again, BBs released in the third quarter are almost exclusively about economic activity in that quarter rather than the previous quarter. Figure 11 shows that during the recession period, 1 October 2007 to 31 December 2009, the BB information released in the third month in the quarter improves MSPE of real GDP growth forecasts growth by over 15% and nearly 10% up 2 weeks after the BB release date. This suggests

[16] In periods where data for $\tilde{y}^{GDP}_{t_Q - D_{GDP,1}}$ is not available, we use the forecast value for this variable implied by the model.
[17] We do not conduct a formal statistical test of forecast difference as the forecast horizon changes at every BB release date, getting progressively shorter as new observations arrive with the quarter and then jumping up as the sample moves into the beginning of the next quarter. The standard tests assume that the forecast horizon is constant and do not take into account a varying forecast horizon.
[18] We examined time period up to 18 days after the BB release. Beyond 18 days there were too few observations in the third month of the quarter.

Copyright © 2016 John Wiley & Sons, Ltd

1099131x, 2017, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/for.2450 by Plymouth University, Wiley Online Library on [08/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

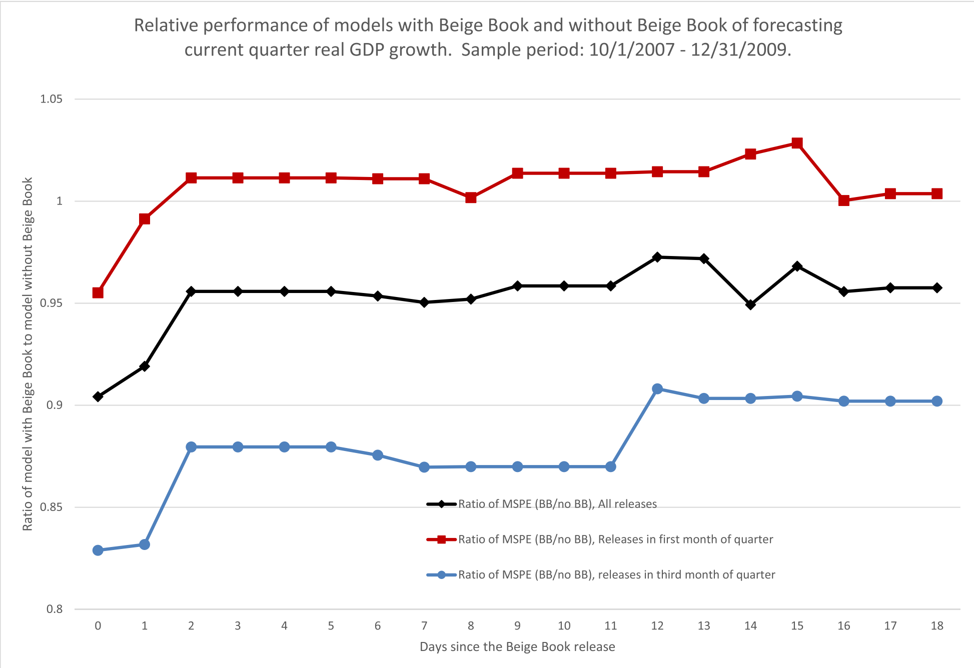

Figure 11. Relative performance of models with and without BB of forecasting current-quarter real GDP growth. Sample period: 1 October 2007 to 31 December 2009. [Colour figure can be viewed at wileyonlinelibrary.com]

that the BB contains potentially useful information about real GDP growth above and beyond that contained in the other ADS indicators, but the informational advantage is more substantial later in the quarter than earlier in the quarter.

CONCLUSION

In this paper, we used text analytics to develop a quantitative index of the Beige Book. This index was in turn included as an observable variable in a dynamic factor model of economic activity based on Aruoba *et al.* (2009). We find that when all updated information about the observables is available the ADS index with and without the BB are very similar, suggesting that including the BB does not alter our historical view of economic activity. On the other hand, we find that at the time the BB is released using the information in the BB can yield modest improvements in the estimated ADS index—moving it towards the full-sample estimates (the 'truth'). This is more likely to be that case around turning points than in more normal times. Unfortunately, the additional information in the BB gets stale fast; by the time the BB is 3 weeks old it provides very little additional information not already included in other more recently released quantitative indicators.

Overall, our results suggest that using the BB information yields modest improvements in 'real time' in our inference about underlying economic activity. The development of computer-based algorithms to 'read' the BB has substantially lowered the costs of including information contained in the BB into quantitative index models of economic activity, so that these models can be continually updated and evaluated. It also opens the possibility that we can systematically learn from the BB.

There are some directions for future research. Even though we used sophisticated text analytics, our computer essentially read the BB 'out-of-context'. That is, aside from using real GDP growth to weight the different BB components, when conducting the text analytics we only used the information in the BB and ignored other information 'outside' the BB. It is conceivable that employing other quantitative information might allow one to better extract the word patterns/combinations that are of the greatest interest to an analyst, just like a human reader uses context to extract meaning from the text.

APPENDIX A

The dynamic factor model in the text can be written in state-space form:

$$\tilde{y}_t = HS_t + AX_t + u_t^* \tag{A1}$$

$$S_t = F_t S_{t-1} + Z\epsilon_t \tag{A2}$$

where equation (A1) is the observation equation and (A2) is the state equation. The vector of observables is given by

$$\tilde{y}_t = \left(\widetilde{\text{JC}}_t,\ \Delta\widetilde{\text{GDP}}_t,\ \widetilde{\text{BB}}_t\ \Delta\widetilde{\text{IP}}_t,\ \Delta\tilde{I}_t,\ \Delta\widetilde{\text{MT}}_t,\ \Delta\widetilde{\text{EM}}_t \right)' \tag{A3}$$

Copyright © 2016 John Wiley & Sons, Ltd

1099131x, 2017, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/for.2450 by Plymouth University, Wiley Online Library on [08/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table A1. Full-sample parameter estimates of ADS model with BB index included

Parameter	Estimate	SE	t-stat
β_{JC}	−0.007	9.31E − 04	−7.24
β_{GDP}	0.010	9.35E − 04	10.9
$\beta_{\text{Beige Book}}$	0.012	1.47E − 03	7.87
β_{IP}	0.032	2.41E − 03	13.1
$\beta_{\text{personal income}}$	0.015	1.76E − 03	8.74
$\beta_{\text{manufacturing}}$	0.026	2.09E − 03	12.3
$\beta_{\text{employment}}$	0.022	1.80E − 03	12.4
γ_{JC}	0.950	5.75E − 03	1.65
γ_{GDP}	−0.059	5.91E − 02	−0.993
$\gamma_{\text{Beige Book}}$	0.340	4.91E − 02	6.92
γ_{IP}	−0.193	4.37E − 02	−4.42
$\gamma_{\text{personal income}}$	−0.199	4.21E − 02	−4.72
$\gamma_{\text{manufacturing}}$	−0.420	3.91E − 02	−10.7
$\gamma_{\text{employment}}$	0.252	4.22E − 02	5.96
σ_{JC}	0.234	3.45E − 03	67.7
σ_{GDP}	0.638	3.78E − 02	16.9
$\sigma_{\text{Beige Book}}$	0.722	2.69E − 02	26.8
σ_{IP}	0.608	2.96E − 02	20.5
$\sigma_{\text{personal income}}$	0.905	2.86E − 02	31.7
$\sigma_{\text{manufacturing}}$	0.745	2.72E − 02	27.3
$\sigma_{\text{employment}}$	0.577	2.14E − 02	27.0
ρ	0.989	1.29E − 03	769.8

and the state variables are

$$S_t = \left(x_t,\ x_{t-1},\ \ldots\ x_{t-p+1},\ C_t^{\text{W}},\ C_t^{\text{Q}},\ C_t^{\text{M}},\ C_t^{\text{BB}} \right)' \tag{A4}$$

The vector of exogenous and predetermined variables in the observation equation is given by

$$\begin{aligned} X_t = (1, &\widetilde{\text{JC}}_{t-\text{W}}, \Delta\widetilde{\text{GDP}}_{t-\text{Q}}, \widetilde{\text{BB}}_{t-\text{BB}}, \Delta\widetilde{\text{IP}}_{t-\text{M}}, \Delta\tilde{I}_{t-\text{M}}, \Delta\widetilde{\text{MT}}_{t-\text{M}}, \\ &\Delta\widetilde{\text{EM}}_{t-\text{M}} \ldots\ldots \widetilde{\text{JC}}_{t-nW}, \Delta\widetilde{\text{GDP}}_{t-n\text{Q}}, \widetilde{\text{BB}}_{t-n\text{BB}}, \Delta\widetilde{\text{IP}}_{t-nM}, \\ &\Delta\tilde{I}_{t-n\text{M}}, \Delta\widetilde{\text{MT}}_{t-n\text{M}}, \Delta\widetilde{\text{EM}}_{t-n\text{M}})' \end{aligned} \tag{A5}$$

Finally, the vector of i.i.d. shocks in the observation equation is given by

$$u_t^* = \left(u_t^{*1},\ u_t^{*2},\ u_t^{*3},\ u_t^{*4},\ u_t^{*5},\ u_t^{*6},\ u_t^{*7} \right)' \tag{A6}$$

Note that $\text{var}(u_t^{*i}) = \sigma_i^2,\ i = 1, 2, \ldots, 7$.

The factor loadings are given in the matrix H, where

$$H = \begin{pmatrix} 0_{1*p} & \beta^{\text{JC}} & 0 & 0 & 0 \\ 0_{1*p} & 0 & \beta^{\text{GDP}} & 0 & 0 \\ 0_{1*p} & 0 & 0 & \beta^{\text{BB}} & 0 \\ 0_{1*p} & 0 & 0 & 0 & \beta^{\text{IP}} \\ 0_{1*p} & 0 & 0 & 0 & \beta^{I} \\ 0_{1*p} & 0 & 0 & 0 & \beta^{\text{MT}} \\ 0_{1*p} & 0 & 0 & 0 & \beta^{\text{EM}} \end{pmatrix} \tag{A7}$$

while the coefficients on the exogenous and predetermined variables are contained in the matrix

$$A = \begin{pmatrix} c_1^* & \gamma_{11} & & & \cdots & \gamma_{1n} & & \\ c_2^* & & \gamma_{21} & 0 & \cdots & & \gamma_{2n} & 0 \\ \vdots & & 0 & \ddots & \cdots & & 0 & \ddots \\ c_7^* & & & \cdots\ \gamma_{71} & \cdots & & & \gamma_{7n} \end{pmatrix} \tag{A8}$$

Copyright © 2016 John Wiley & Sons, Ltd

1099131x, 2017, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/for.2450 by Plymouth University, Wiley Online Library on [08/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The autoregressive parameters in the state equation are time varying and are contained in the matrix

$$F_t = \begin{pmatrix} \rho_1 & \cdots & \rho_{p-1} & \rho_p & 0 & 0 & 0 & 0 \\ 1 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 \\ & \ddots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 & 0 & 0 & 0 \\ \rho_1 & \cdots & \rho_{p-1} & \rho_p & \zeta_t^{\mathrm{W}} & 0 & 0 & 0 \\ \rho_1 & \cdots & \rho_{p-1} & \rho_p & 0 & \zeta_t^{\mathrm{Q}} & 0 & 0 \\ \rho_1 & \cdots & \rho_{p-1} & \rho_p & 0 & 0 & \zeta_t^{\mathrm{BB}} & 0 \\ \rho_1 & \cdots & \rho_{p-1} & \rho_p & 0 & 0 & 0 & \zeta_t^{\mathrm{M}} \end{pmatrix} \tag{A9}$$

For the purposes of filtering, we assume the ζ_t^i is deterministic and known. This assumption is reasonable as the first day of the relevant period is known in advance. Finally, the loading matrix of the shock to the state vector

$$Z = (1,\ 0,\ \ldots,\ 0,\ 1,\ 1,\ 1,\ 1)' \tag{A10}$$

For the estimated model, the number of lagged observables was set to one ($n = 1$), the length of the autoregressive process for x_t was set to one ($p = 1$) and $\mathrm{var}(\epsilon_t) = 1 - \rho^2$.

ACKNOWLEDGEMENTS

We thank Tom Fomby, Olivier Coibion and participants at conferences at the Bank of Korea, Federal Reserve Day-before Regional System Meetings, and Southern Economic Association for helpful comments. The views expressed in this paper do not necessarily reflect those of the Federal Reserve Bank of Dallas or the Federal Reserve System.

REFERENCES

Armesto M, Hernandez-Murillo R, Owyang MT, Piger JM. 2009. Measuring the information content of the Beige Book: a mixed data sampling approach. *Journal of Money, Credit, and Banking* **41**(1): 35–55.

Aruoba SB, Diebold FX, Scotti C. 2009. Real-time measurement of business conditions. *Journal of Business and Economic Statistics* **27**(4): 417–427.

Balke NS, Petersen D'A. 1998. How well does the Beige Book reflect economic activity? Evaluation qualitative information quantitatively. *Federal Reserve Bank of Dallas Working Paper*: 98–02.

Balke NS, Petersen D'A. 2002. How well does the Beige Book reflect economic activity? Evaluating qualitative information quantitatively. *Journal of Money, Credit, and Banking* **34**: 114–136.

Balke NS, Yucel M. 2000. Evaluating the Eleventh District's Beige Book. *Federal Reserve Bank of Dallas, Economic and Financial Review* **2000**: 2–10.

Berge TJ, Jorda O. 2011. Evaluating the classification of economic activity into recessions and expansions. *American Economic Journal: Macroeconomics* **3**: 246–277.

Blinder A. 1997. Distinguished lecture on economics in government: what central bankers could learn from academics—and vice versa. *Journal of Economic Perspectives* **11**(Spring): 3–19.

Chauvet M, Piger J. 2008. A comparison of the real-time performance of business cycle dating methods. *Journal of Business and Economic Statistics* **26**: 42–49.

Clark TE, West KD. 2007. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics* **138**: 291–311.

Fettig D, Rolnick AJ, Runkle DE. 1999. The Federal Reserve's Beige Book: a better mirror than crystal ball. *Federal Reserve Bank of Minneapolis: The Region* **1999**: 28–32.

Fulmer M. 2014. *A Text Analytics analysis of the Federal Reserve Beige Book Corpus*. Southern Methodist University: Doctoral dissertation.

Gatu C. 2006. Branch-and-bound algorithms for computing the best-subset regression models. *Journal of Computational and Graphical Statistics* **15**: 139–156.

Ginther DK, Zavodny M. 2001. The Sixth District Beige Book: timely information on the regional economy. *Federal Reserve Bank of Atlanta: Economic Review* **86**: 19–29.

Giusto A, Piger J. 2014. Identifying the business cycle turning points in real time with vector quantization. *International Journal of Forecasting, forthcoming 2017*.

Harvey AC. 1989. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press: Cambridge, UK.

Miller A. 2002. *Subset Selection in Regression* (2nd edn, ed.) Chapman & Hall: London.

Payne DR. 2001. Anticipating monetary policy with the Federal Reserve's Beige Book: Re-specifying the Taylor rule. *Business Economics* **36**: 21–30.

Sadique S, In F, Veeraraghavan M, Wachtel P. 2013. Soft information and economic activity: Evidence from the beige book. *Journal of Macroeconomics* **73**: 81–92.

Zavodny M, Ginther DK. 2005. Does the Beige Book move financial markets? *Southern Economic Journal* **78**(1): 138–151.

Copyright © 2016 John Wiley & Sons, Ltd

1099131x, 2017, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/for.2450 by Plymouth University, Wiley Online Library on [08/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Authors' biographies:

Nathan S. Balke is the 2007 Dedman Family Distinguished Professor in Economics at Southern Methodist University. His research interests include macroeconomics and applied time series.

Michael Fulmer receive his Ph.D. in Economics from Southern Methodist University. He is currently with Google, Inc. His research interests include statistics and data analytics.

Ren Zhang is a Ph.D. student in Economics at Southern Methodist University. He research interests include macroeconomics and time series analysis.

Authors' addresses:
Nathan S. Balke, Department of Economics, Southern Methodist University, Dallas, TX 75275-0496, USA; and Federal Reserve Bank of Dallas, 2200 N Pearl St, Dallas, TX 75201 USA.
Michael Fulmer and **Ren Zhang**, Department of Economics, Southern Methodist University, Dallas, TX 75275-0496, USA.

Copyright © 2016 John Wiley & Sons, Ltd

1099131x, 2017, 5, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/for.2450 by Plymouth University, Wiley Online Library on [08/04/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License