JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. (2015)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/jae.2494

FORECASTING CONSUMPTION: THE ROLE OF CONSUMER CONFIDENCE IN REAL TIME WITH MANY PREDICTORS

KAJAL LAHIRI,[a*] GEORGE MONOKROUSSOS[b] AND YONGCHEN ZHAO[c]

[a] *Department of Economics, University at Albany, SUNY, NY, USA*
[b] *European Commission, Joint Research Centre, Ispra, Italy*
[c] *Department of Economics, Towson University, Towson, MD, USA*

SUMMARY

We study the role of consumer confidence in forecasting real personal consumption expenditure, and contribute to the extant literature in three substantive ways. First, we re-examine existing empirical models of consumption and consumer confidence, not only at the quarterly frequency, but using monthly data as well. Second, we employ real-time data in addition to commonly used revised vintages. Third, we investigate the role of consumer confidence in a rich information context. We produce forecasts of consumption expenditures with and without consumer confidence measures using a dynamic factor model and a large, real-time, jagged-edge dataset. In a robust way, we establish the important role of confidence surveys in improving the accuracy of consumption forecasts, manifesting primarily through the services component. During the recession of 2007–2009, sentiment is found to have a more pervasive effect on all components of aggregate consumption: durables, non-durables and services.

Received 15 April 2013; Revised 14 July 2015

1. INTRODUCTION

The concept of *animal spirits*, in the standard Keynesian sense, has influenced economic thinking for a long time and has received renewed and intense attention in the run-up to the recent financial crisis and the ensuing recession (cf. Akerlof and Shiller 2010). The confidence of economic agents and their importance to the economy occupies a central role in this discussion. Consumer confidence, in particular, is typically at the center of attention of the business press. It has also been studied extensively by academics as well as policymakers.[1] This interest is certainly justified given the presumed role of confidence in shaping aggregate consumption spending.

Not surprisingly, many academic studies, both for the USA and internationally, investigate various aspects of the relationship between consumer confidence and consumer spending, at both the micro and the aggregate levels. Souleles (2004) found that aggregate shocks do not hit all segments of the population equally; rather they are systematically mediated by demographic characteristics of households. In addition, given the timing advantage of the standard measures of confidence, as was first emphasized by Howrey (2001), consumer confidence may have important implications for monitoring the economy in real time and for economic policy, as well as for testing key economic theories, such as the canonical permanent income–rational expectations (PI/RE) hypothesis.

Consequently, a central preoccupation of the relevant literature, including the papers cited above, is to assess the forecasting power of consumer confidence for consumer spending at the aggregate level. As discussed by Ludvigson (2004), some evidence to that effect is generally found in the literature, but it becomes more modest once a few additional variables that have traditionally been considered in studies of consumer confidence are added to the specification. However, much of this existing literature has several limitations.

* Correspondence to: Kajal Lahiri, Department of Economics, University at Albany, SUNY, 1400 Washington Avenue, Albany, NY 12222, USA. E-mail: klahiri@albany.edu

[1] For a worldwide review and assessment of consumer sentiment surveys, see Curtin (2007) and references therein.

First, quarterly data are commonly used. However, the most widely known measures of consumer confidence (the University of Michigan's Index of Consumer Sentiment (ICS) and the Conference Board's Consumer Confidence Index (CCI)) are available at a monthly frequency, and employing quarterly averages of these monthly indices in models of consumption expenditures may conflate the monthly effects of consumer confidence.[2] Moreover, consumer spending itself, and also many other relevant indicators, are available on a monthly basis as well.

Second, revised data on the relevant variables are employed, as opposed to the data that were actually available in real time, i.e. before any revision that only became available at subsequent points in time. Of course, for monetary policy purposes or, more generally, for the purpose of assessing the real-time forecasting power of consumer confidence, real-time data should be used.

Third, the regression models used to assess the predictive power of consumer confidence typically include only a small number of additional variables, i.e. a rather small information set, whereas many more variables, possibly in the hundreds, are available that are potentially relevant to consumption decisions.

In this study, we provide what is arguably a more realistic assessment of the predictive power of consumer confidence on consumer spending by addressing all of the three issues mentioned above using more recent data.

Our starting point is Ludvigson (2004). We extend some of the existing models using monthly and real-time data, in addition to quarterly and revised vintages. We then employ a large real-time dataset with close to 200 explanatory variables at the monthly frequency in order to assess the marginal impact of confidence on consumer spending in the context of such a large information set in real time. In this setting, a dynamic factor model is preferred to deal with the challenges, such as the proliferation of parameters (Stock and Watson, 2011; Banbura *et al.*, 2013). Through a series of exercises using the framework first developed by Giannone *et al.* (2008), we gain insight into the marginal impact of consumer confidence on consumer spending in real time by comparing consumption forecasts based on information sets with and without consumer confidence measures. In contrast to much of the existing literature, we consider both in-sample and out-of-sample forecasts. Our results generally establish the undeniable, though modest, importance of consumer confidence in forecasting aggregate consumption.

The rest of the paper is organized as follows. Section 2 provides some discussion of the important aspects of the consumption and consumer confidence data. In Section 3, we first revisit some models used in the existing literature on the predictive power of consumer confidence. Then, we outline the dynamic factor approach and assess the predictive power of consumer confidence in real time when it is a part of a large information set. We also provide a detailed discussion of our findings in Section 3. Section 4 concludes.

2. CONSUMPTION AND CONSUMER SENTIMENT: A CLOSER LOOK AT THE DATA

Consumer spending accounts for about two-thirds of domestic final spending in the USA. The primary measure of consumer spending on various types of goods and services is real personal consumption expenditure (PCE). It covers purchases made by households and nonprofit institutions serving households (NPISHs). PCE data come from Personal Income and Outlays released by the Department of Commerce, Bureau of Economic Analysis (BEA). It can be measured by type of products or by function (health, recreation, communication, etc.). In this study we examine the total PCE and PCE by main types of products: durable goods, non-durable goods, and services.

PCE data are available at both monthly and quarterly frequencies. The quarterly series are released every month together with the GDP series, in the last week of the month. Similar to the GDP series, there is a one-quarter lag between the end of a period and the release of data covering that period.

[2] In what follows, we use the words 'sentiment' and 'confidence' interchangeably.

DOI: 10.1002/jae

Advance estimates of PCE are released for the previous quarter at the end of the first month of each quarter. At the end of the second and the third months, the preliminary and the final estimates for the previous quarter are released, respectively. The monthly series are released 1 day after the release of the quarterly series. The publication lag for the monthly series is 1 month. Monthly PCE series are also subject to revisions. Such revisions are announced in the subsequent monthly releases.

The monthly PCE values for the first 2 months (released at the end of the second and the third month of a quarter) play an important role in forecasting the quarterly PCE for that quarter and beyond, before the advance release of the quarterly value becomes available. The advance release and the first two monthly values do not pin down the third monthly value because of data revisions. This implies that to someone forecasting in real time, every monthly release, in addition to the advance quarterly announcement, contains some additional information not present in any of the previous releases. To our knowledge, monthly PCE series have not been used in any study of the relationship between consumer confidence and consumer spending. We thus consider using the monthly consumption series (in addition to their quarterly counterparts) as one of our main contributions.

One of the most recognized measures of consumer confidence is the Index of Consumer Sentiment (ICS) from the Survey of Consumers administered by the Survey Research Center of the University of Michigan (CAB). The Survey started as an annual survey in 1946 (cf. (Katona, 1951)). It became a quarterly survey in 1952 and then a monthly survey in 1976. The ICS index can be separated into a present conditions index and an expectations index, based on the questions used in constructing the index. Each month, about 500 households are interviewed by phone. The preliminary releases of the index come out around mid-month, based on the information gathered in the first half of the month, usually two-thirds of the full sample. The final releases are scheduled on the last Friday of each month. Public and media attention is usually concentrated on the final releases, which are quite timely, and subject to no further revision.

The Survey of Consumers tracks many different aspects of consumer attitudes and expectations. About 50 core questions are asked in each survey. Five of these questions are used to construct the ICS, and they are as follows:

(i) *We are interested in how people are getting along financially these days. Would you say that you (and your family living there) are better off or worse off financially than you were a year ago?*
(ii) *Now looking ahead—do you think that a year from now you (and your family living there) will be better off financially, or worse off, or just about the same as now?*
(iii) *Now turning to business conditions in the country as a whole—do you think that during the next twelve months we'll have good times financially, or bad times, or what?*
(iv) *Looking ahead, which would you say is more likely—that in the country as a whole we'll have continuous good times during the next five years or so, or that we will have periods of widespread unemployment or depression, or what?*
(v) *About the big things people buy for their homes—such as furniture, a refrigerator, stove, television, and things like that. Generally speaking, do you think now is a good or bad time for people to buy major household items?*

Among these questions, (i) and (v) are mainly about the present conditions of the household and the economy, whereas the remaining three questions are clearly about household expectations. For each of the five questions, a respondent can choose among three responses: favorable (e.g. situation getting better), neutral (e.g. situation is the same as before), and unfavorable (e.g. situation getting worse). The relative score for each question, known as diffusion index, is calculated as the percentage of respondents giving favorable replies minus the percentage giving unfavorable replies plus 100. Scores are then rounded, added up and divided by the base period value of 1966 to form an index. Based on the content of the questions, the Index of Consumer Expectations (ICE) is constructed using

DOI: 10.1002/jae

relative scores for questions (ii)–(iv). Similarly, the Index of Current Economic Conditions (ICC) is constructed using relative scores for questions (i) and (v). Finally, the overall index, ICS, is constructed using all five relative scores.

Another widely used measure of consumer confidence is the Consumer Confidence Index (CCI) from the Consumer Confidence Survey administered by the Conference Board (TCB). The survey began in 1967 as a bi-monthly survey. Since June 1977, the survey has been administered monthly. Similar to the ICS, the CCI can also be separated into two components: the present situation component and the expectations component. Each month, a mail survey is sent out and approximately 3000 completed questionnaires are collected.[3] Preliminary estimates are based on survey responses collected before the 18th of each month. Final estimates are published with the release of the following month's data, scheduled on the last Tuesday of each month.

The CCI and its two components are also based on five questions. The first two are used to construct the present situations index and the rest are used to construct the expectations index. All five questions are used to construct the CCI. These questions are:

(i) *How would you rate the present general business conditions in your area?*
(ii) *What would you say about available jobs in your area right now?*
(iii) *Six months from now, do you think general business conditions will be better, the same, or worse?*
(iv) *Six months from now, do you think there will be more, the same, or fewer jobs available in your area?*
(v) *How would you guess your total family income to be six months from now? Answers: higher, the same, or lower?*

To each of the five questions, three response options are available: positive, neutral, or negative. The proportion of respondents giving positive responses among those who do not give neutral responses for each question is computed first. A corresponding index is produced for each proportion with the average value for all months in 1985 as the benchmark. Finally, relevant indices are averaged to produce the CCI and its two components. Seasonal adjustment is performed where needed.

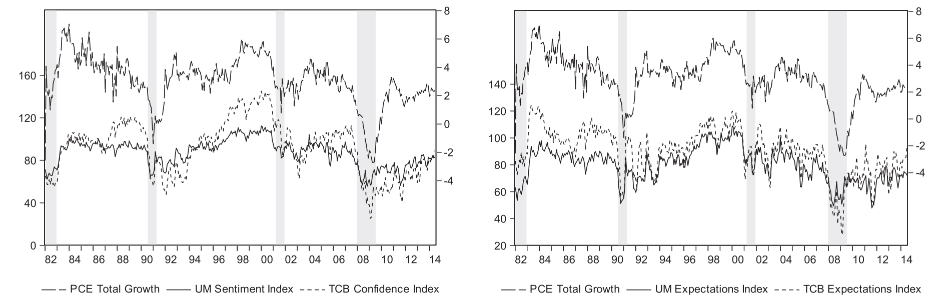
In this study, we focus on consumption and consumer confidence data between January 1982 and June 2014—a total of 390 months (130 quarters).[4] Figure 1 compares the University of Michigan's consumer sentiment and the Conference Board's consumer confidence measures with the 12-month moving average of annualized (advanced estimates of) monthly growth in real personal consumption expenditure. The shaded areas are periods of NBER-defined recessions. Figure 1(a) shows the two overall indices and Figure 1(b) shows the two expectations components. It is clear from the figure that over business cycles the confidence measures evolve in a similar way to consumption growth. We did not find a clear lead or lag relationship between the two expectations components. This is an interesting observation, since one of the Michigan survey questions concern up to 5 years in the future, whereas the Conference Board survey questions cover only up to 6 months. As shown in the experiments that follow, this difference in the horizon of the questions does not lead to a systematic difference in the forecasting power of the two measures. We also observe from Figure 1 that the two measures themselves are not always that closely correlated with each other, and the variance of CCI is almost twice that of ICS primarily due to the methods of construction.

---

[3] A new sample design was introduced effective November 2010. A discussion of historical comparability is available in the Consumer Confidence Survey Technical Note (http://www.conference-board.org/pdf_free/press/TechnicalPDF_4134_1298367128.pdf).

[4] While it is possible to estimate the models in Section 3.1 using data from January 1978, for the sake of consistency across exercises we start from 1982 owing to the unavailability of the large real-time dataset used in the rest of Section 3. Our conclusions stay the same if we used the longer sample in Section 3.1.

DOI: 10.1002/jae



(a) Overall indices **(b)** Expectations indices

Figure 1. Sentiment measures and total consumption growth. This figure shows the overall index (left) and the expectations index (right) of both the University of Michigan (UM, solid line) and the Conference Board (TCB, short dashed line) measure (both on left axis), compared with 12-month moving average of annualized monthly percentage growth in real total personal consumption expenditure (PCE total, long dashed line, on right axis)

As mentioned above, Figure 1 plots real-time data of consumption. The importance of using real-time data for tasks such as assessing monetary policy or evaluating forecasts in general cannot be overemphasized (see, among others, Croushore-and-Stark 2001, Orphanides 2001). More specifically in our context, and as Ludvigson (2004) discusses, it is essential not to use currently available revised data when assessing the real-time forecasting power of consumer confidence for consumption. However, much of the existing literature on consumer confidence and consumer spending has used revised data. Therefore, in the estimation and forecast evaluation exercises that follow, we employ both current/latest and first vintages. Specifically, in all our real-time exercises, the first vintage of consumption is used to evaluate forecasts, while our estimation uses appropriate vintages specified in the description of each empirical exercise below.

Furthermore, and as discussed in the Introduction, the existing literature employs quarterly data, which can mask important information available at monthly frequency. Thus, in the exercises that follow, we use both quarterly and monthly data.

3. FORECASTING CONSUMPTION: EXISTING MODELS AND THE FACTOR APPROACH

3.1. Existing Models Revisited

A number of attempts have been made in the literature to quantify the importance of consumer confidence in explaining and predicting quarterly consumption expenditures. In this subsection, we first re-examine the main empirical models used in previous studies (e.g. Carroll *et al.*, 1994, Bram-and-Ludvigson 1998, Ludvigson 2004), and then extend them to model monthly consumption expenditures. In the process, we focus on the significance of the confidence measure and the change in a model's explanatory power due to the addition of this measure. As discussed in more detail later in the section, the models examined in this subsection are rather restrictive in terms of the additional information they use, even though they have been the workhorse in the literature. Owing to their importance in the macro literature, we attempt to extend them in two important ways in this subsection: (i) we use both revised and real-time data to estimate these models; and (ii) in addition to modeling quarterly consumption, we model monthly consumption as well.

To investigate whether sentiment measures contain unique information that is not available in other aggregate measures of economic activity, we consider a simple model of consumption expenditure. For time period t, let C_t be (a type of) consumption expenditure, and let S_t be (a measure of) consumer confidence. We estimate the following model:

DOI: 10.1002/jae

$$\Delta \ln(C_t) = \alpha_0 + \sum_{i=1}^{\tau} \alpha_i \cdot Z_{t-i} + \sum_{i=1}^{\tau} \alpha_i \cdot \Delta \ln(C_{t-i}) + \sum_{i=1}^{\tau} \beta_i \cdot S_{t-i} + \varepsilon_t \tag{1}$$

where, following the literature, the number of lags τ is set to 4. Apart from lagged values of consumption expenditures, there is a set of baseline macroeconomic variables Z_{t-i} that are typically included in the existing literature (see, for instance, Carroll *et al.*, 1994, Ludvigson 2004). They include the return to S&P 500 index, the 3-month Treasury Bill rate, and labor income growth, which is wages and salaries plus transfers minus personal contributions for social insurance.

The consumption expenditure C_t is one of the following four: total personal consumption expenditure (Total), expenditure on durable goods (Durable), expenditure on non-durable goods (Non-Durable), and that on services (Services). The confidence measure is one of the following four: the expectations components and the overall indices from the University of Michigan and the Conference Board.[5] Quarterly sentiment measures are averages of all monthly values within a quarter. Standard error estimates are robust to heteroskedasticity and serial correlation. We estimate the quarterly models using both real-time and revised data. But the monthly models are estimated using revised series only.[6]

Table I presents the estimation results. In general, the model's explanatory power is similar to that found in the existing literature using quarterly data (see Ludvigson, 2004). In the table, incremental $\bar{R}^2$s are the difference between the $\bar{R}^2$ of the models with and without the S_{t-i} terms in equation (1). The joint significance of coefficients of all lags of the confidence measure is also reported. The table shows that the consumer confidence measures do indeed explain consumption, but the magnitude varies. Inclusion of consumer confidence increases the $\bar{R}^2$ in all 48 models.[7] On average, in quarterly models using real-time data and the University of Michigan sentiment measures, a 2.8% increase in $\bar{R}^2$ is observed. In models using the Conference Board measures, a 5.4% increase in $\bar{R}^2$ is observed. For the monthly models using revised data, in models using the University of Michigan sentiment measures, a 2.9% increase in $\bar{R}^2$ is observed on average. In models using the Conference Board confidence measures, an average increase of 3.0% is observed. With real-time data, sentiment significantly affects services and non-durable goods consumption with similar magnitudes. These two components of consumption have very similar behavior over our sample.[8]

These results suggest that the contribution of consumer confidence measures in explaining consumption expenditures is statistically significant in many cases, but is, arguably, of modest size, as found in most other studies.[9] One exception is Barsky and Sims (2012), who find that innovations in sentiment are prognostic of long-run movements in output growth and consumption. They argue that confidence innovations partly reflect shifts in people's beliefs about future productivity growth.

However, while the explanatory variables used here are the standard choices in the macro literature, there are many other variables with potentially significant explanatory power for consumption. It is

[5] In line with the literature, results on the current conditions components of both confidence measures are omitted.

[6] Real-time data for the components of labor income are only available at a quarterly frequency; thus our monthly models are estimated using only the latest vintage. Nominal labor income and the S&P 500 index are converted to real terms using the PCE price index.

[7] There are two data frequencies, four types of consumption, and two measures of confidence for both the Conference Board and University of Michigan. Also the quarterly models are examined using both the real-time data and the revised data.

[8] We conducted the exercises without Z_t as well. Results show that gains from using sentiment are larger in regression models without any additional explanatory variables.

[9] Croushore (2005) finds that confidence is not significant in forecasting consumption in a specification that includes explanatory variables similar to what we used in this subsection. There are several important differences between his exercises and the ones presented here. First, for evaluating forecasts, he uses real-time consumption data, but vintages 'available just prior to a benchmark revision'—these benchmark revisions are made about every 5 years. One may say that evaluating forecasts generated in real time using figures that only became available after several revisions may be a very stringent test that confidence measures are unlikely to pass. Second, whereas our sample covers 1982:Q1 to 2014:Q2, Croushore's sample is from 1992:Q1 to 2002:Q4, which excludes the latest recession.

DOI: 10.1002/jae

Table I. Incremental explanatory power of sentiment/confidence measures

PCE and component	Sentiment	Monthly model (revised data) Inc. $\bar{R}^2$	Relative MSE	p	Quarterly model (revised data) Inc. $\bar{R}^2$	Relative MSE	p	Quarterly model (real-time data) Inc. $\bar{R}^2$	Relative MSE	p
Durable goods	TCB Exp.	0.027	0.956	**0.004**	0.022	0.940	0.272	0.007	0.956	0.307
	TCB Index	0.010	0.978	**0.058**	0.030	0.930	0.163	0.047	0.913	**0.097**
	UM Exp.	0.018	0.968	**0.007**	0.008	0.955	0.338	0.007	0.956	0.506
	UM Index	0.013	0.973	**0.021**	0.013	0.949	0.346	0.018	0.944	0.332
Non-durable goods	TCB Exp.	0.043	0.938	**0.003**	0.020	0.943	0.301	0.124	0.830	**0.030**
	TCB Index	0.014	0.973	**0.051**	0.017	0.946	0.357	0.110	0.845	**0.053**
	UM Exp.	0.033	0.950	**0.002**	0.006	0.957	0.236	0.049	0.911	**0.089**
	UM Index	0.033	0.950	**0.001**	0.006	0.957	0.348	0.062	0.897	**0.089**
Services	TCB Exp.	0.038	0.943	**0.000**	0.056	0.861	**0.000**	0.047	0.901	**0.004**
	TCB Index	0.006	0.982	0.152	0.025	0.918	0.101	0.031	0.922	**0.013**
	UM Exp.	0.023	0.962	**0.004**	0.038	0.894	**0.010**	0.023	0.933	**0.095**
	UM Index	0.024	0.960	**0.005**	0.048	0.875	**0.008**	0.028	0.926	**0.041**
Total	TCB Exp.	0.074	0.905	**0.000**	0.012	0.945	0.221	0.029	0.924	0.285
	TCB Index	0.017	0.970	**0.035**	0.012	0.944	0.296	0.037	0.914	0.286
	UM Exp.	0.051	0.932	**0.000**	0.002	0.959	0.333	0.017	0.941	0.420
	UM Index	0.042	0.942	**0.000**	0.005	0.956	0.365	0.021	0.935	0.436

Note: This table gives the incremental $\bar{R}^2$ (Inc. $\bar{R}^2$; positive means adding sentiment measure increases explanatory power), the relative MSE of the model with/without sentiment measure, and the p-value of the test of the joint significance (p) of four lags of sentiment measure. Newey-West standard errors are used with 4 lags. The p-values in bold are those smaller than 0.1. Dependent variable is $[\ln(\text{PCE}_t) - \ln(\text{PCE}_{t-1})] \times 400$ for quarterly models and $[\ln(\text{PCE}_t) - \ln(\text{PCE}_{t-1})] \times 1200$ for monthly models. The sample covers January 1982 to June 2014.

conceivable that the information contained in consumer confidence measures could simply be a combination of the information found in a large number of macroeconomic indicators not included in the above three variables.

3.2. Dynamic Factor Framework and a Large Dataset

Given the above discussion, we want to explore the role of confidence in forecasting consumption when the forecasts are generated using a wide information set in addition to sentiment. This requires using a large number of additional explanatory variables. One immediate problem that arises with this type of exercise is the lack of degrees of freedom. Even when we use monthly consumption data, the number of available observations is limited to the hundreds. Yet potentially useful variables also come in the hundreds, which of course makes the ordinary least squares (OLS) regression model impractical. So we employ here an approach that allows us to address this challenge.

The approach we use in the following subsections is based on the dynamic factor model of Giannone *et al.* (2008, henceforth GRS). In this subsection we introduce the model and the explanatory variables, and discuss associated issues. In Subsection 3.3 we explore the effect of sentiment on consumption through out-of-sample pseudo-real-time exercises. In Subsection 3.4 we assess the marginal contribution of sentiment to consumption forecasts again, but in the context of real-time data. The GRS framework is particularly suitable for teasing out the marginal effects of specific data releases that are announced regularly at certain times of the month (Banbura *et al.*, 2013). This approach utilizes a dynamic factor model in a state-space form to summarize the common information from a large number of explanatory variables with potentially mixed frequencies and varying patterns of missing data.

DOI: 10.1002/jae

Let x_t be a $N \times 1$ vector of observed independent variables for time period $t = \{1, \ldots, T\}$, and let F_t be a $r \times 1$ vector of latent factors representing the state of the economy. The latent factors drive both the concurrent evolution of the explanatory variables and the future evolutions of the latent factors themselves. This relationship is summarized in a state-space model as follows:

$$x_t = \mu + \Lambda \cdot F_t + \xi_t \tag{2}$$

$$F_t = A \cdot F_{t-1} + B \cdot u_t \tag{3}$$

where ξ_t is an $N \times 1$ vector of variable-specific innovations, Λ is an $N \times r$ matrix of factor loadings, A is an $r \times r$ matrix with all roots of $\det(I_r - Az)$ outside the unit circle, B is an $r \times q$ matrix of rank q, and u_t is a $q \times 1$ vector of common shocks. As is standard in the literature, we set $r = q = 2$.[10]

Given the 'jagged-edge' nature of the data, i.e. the varying missing data patterns in the large number of explanatory variables, especially toward the end of the sample period, estimation of the model is performed in two steps. In the first step, a fully balanced panel of the explanatory variables is created by discarding any observation toward the end of the sample period for which at least one variable is not observed. This is used to obtain preliminary estimates of the latent factors by principal components. These estimates are in turn used to estimate the parameters of the model. Given these estimates, in the second step the Kalman smoother is used to compute the latent factors for the entire sample period, including those periods discarded in the first step. In this process, the Kalman smoother forecasts the latent factors for periods when the observations for certain variables are unavailable. In these calculations, sentiment measures were not included. Our variable of interest, a particular category of consumption expenditure, is assumed to be determined by the latent factors and a lagged measure of consumer confidence. A simple OLS regression can be used to establish the link between consumption expenditure and these predictors, and to forecast future consumption expenditures given consumer confidence and forecasts of the factors. To avoid the need to forecast consumer confidence, the latest available values of sentiment are used in multi-period forecasts.[11]

This two-step procedure is necessary to deal with the jagged-edge data structure, which is caused by varying data release schedules and publication lags across all the explanatory variables. Such a data structure is unavoidable if no information is to be discarded when forecasting. For example, at the end of each month, the variables with a publication lag of 1 month will have one more observation in the dataset than the variables with a publication lag of 2 months.

With the publication lag affecting the structure of the dataset at any given time, there are three main determinants of the value of an explanatory variable in this context. The first determinant is the information content of the variable. If it contains only information that comes from other variables, in the sense that it is highly collinear with those variables, then the addition of this variable to the specification will not affect the forecasts made using these dynamic factors alone. The second determinant is the timeliness of the release of the variable. The shorter the publication lag, or the earlier its release date is within a month, the more useful the variable is likely to be. The last determinant is data revisions. Even though its effect on forecast accuracy is unclear, there should be no doubt that the role a variable plays in forecasting in real time cannot be fully ascertained without considering its effect using real-time data before data revisions.

[10] Two principal component factors explain 78% of the variations. The gain from a third factor is relatively small.

[11] This implies that the specification employed is the following: $\Delta \ln(C_t) = \alpha + D \cdot F_t + \beta \cdot S_{t-h} + v_t$. In the exercises below we compare h-step-ahead forecasts of consumption ($\ln(\hat{C}_{t+h})$) made 'with' and 'without' the latest observation (S_t) of sentiment. Forecasts made 'with' the latest observation of sentiment are based on S_t, and forecasts made 'without' the latest observation of sentiment are based on S_{t-1}. We also compare forecasts made 'with' and 'without' the sentiment variable, where the former are based on S_t, and the latter are not based on any sentiment measure, i.e. made only using the factors F_t.

DOI: 10.1002/jae

Corresponding to the three determinants above, we conduct three exercises. In the first exercise, we examine the in-sample fit of the model with and without the entire series of a confidence measure. Here we focus on the first determinant: the information content. This exercise uses only one dataset representing the latest available information/vintage as of June 2014 without any jagged edges. The dataset contains 163 variables and 390 observations from January 1982 to June 2014.

In the second exercise, which is a pseudo-real-time exercise, we attempt to reconstruct a series of 'snapshots' of the jagged-edge data based on a stylized calendar of data release schedules and publication lags, but still using only revised/latest vintage data again.[12] We examine the accuracy of consumption forecasts made with and without the latest release of a confidence measure, while all the historical values of this measure are always in the dataset. Our estimation sample starts from January 1982, as in the previous exercise. Our evaluation sample starts from January 1995 to allow for a sufficiently large estimation sample in making the first pseudo-real-time forecast. The design of this exercise allows us to assess the value of consumer confidence measures attributable to the timeliness of their release. In particular, this assessment is conducted while controlling for any effect of confidence measures that may be due to data revisions to other variables in the large dataset, including that in consumption measures.

In the last exercise, we construct real-time datasets as they were actually available to forecasters in the past. More specifically, we painstakingly recreate a real-time dataset for every Tuesday and every Friday (corresponding to the release date of the two confidence measures) from March 2005 to June 2014. Each dataset contains the appropriate data vintages with a jagged edge reflecting data release schedules.[13] We use these datasets to examine the role of consumer confidence measures in forecasting consumption in real time.[14] In addition to the two determinants of the value of sentiment considered above in exercises 1 and 2 (viz. information content and timeliness of data release), this exercise accounts for the effect of data revisions on forecast accuracy. As before, our estimation sample starts from January 1982.

We separately conduct the three exercises using monthly and quarterly consumption data. Table II summarizes the features of these exercises.

Most of our explanatory variables and their real-time data vintages come from the dataset put together by GRS. This dataset consists of tens of macro variables for the US economy starting in January of 1982. These variables, most of which are at the monthly frequency, include real and monetary quantities, prices, and surveys. Variables used in the exercises in Section 3.1 were added to their dataset. The University of Michigan sentiment indices, consumption expenditures, real GDP, and a few proprietary series are removed from the set of independent variables, based on which the factors are generated. There are 163 variables in the resulting dataset. We obtained real-time data for the variables used in Section 3.1 and personal consumption expenditure and its components from the Federal Reserve Bank of Philadelphia's Real-Time Data Research Center. For quarterly consumption expenditures, monthly vintages are used instead of quarterly vintages, since we produce forecasts of quarterly consumption expenditures at a monthly frequency.

---

[12] All the details on the schedule of data releases and publication lags are available in the GRS paper and at the author's website (http://homepages.ulb.ac.be/~dgiannon).

[13] The real-time dataset, updated every week, was kindly provided to us by David Small of the Federal Reserve Board. This is the same dataset used by Giannone *et al*. (2010), except for a few proprietary series. We made a significant amount of adjustments to these data based on schedules of data releases and revisions in order to recreate the real-time information sets for every Tuesday and every Friday. For example, and in most cases, about 32 variables are updated (either with a new release or with a revision) between a Tuesday and the following Friday. Examples of variables that are typically released during this time include personal income, consumption, new residential sales, inventory, and new orders.

[14] Note that for the Conference Board confidence series the final release for any given month first becomes available with the following month's preliminary release (unlike the University of Michigan measures, which are always available during the current month). To be consistent with the real-time nature of this exercise, we use the preliminary Conference Board data. To our knowledge, we are the first to employ these preliminary series, which were generously made available to us by TCB.

DOI: 10.1002/jae

Table II. Overview of empirical exercises using the factor model with many predictors

Exercise	Table	Jagged-edge data	Vintage of variables (same for both dependent variables and independent variables)	Sentiment variable used in constructing forecasts with sentiment	Sentiment variable used in constructing forecasts without sentiment	Estimation sample	Evaluation sample[e]
Exercise 1: In-sample exercise	3	No[a]	Latest vintage	All sentiment values up to June 2014	None[c]	Jan. 1982 to Jun. 2014	Jan. 1982 to Jun. 2014
Exercise 2: Out-of-sample/pseudo real-time exercise	4	Yes	Latest vintage	All sentiment values up to the end of the estimation sample	Sentiment values up to one period before the end of the estimation sample	Jan. 1982 to Dec. 1994	Jan. 1995 to Jun. 2014
Exercise 3: Real-time exercise (w/wo latest sentiment value)	5	Yes	Historical vintages[b]	All sentiment values up to the end of the estimation sample	Sentiment values up to one period before the end of the estimation sample	Jan. 1982 to Mar. 2005	Mar. 2005 to Jun. 2014
Exercise 3: Real-time exercise (w/wo entire sentiment variable)	6	Yes	Historical vintages[b]	All sentiment values up to the end of the estimation sample	None	Jan. 1982 to Mar. 2005[d]	Mar. 2005 to Jun. 2014

Note: This table gives the specifics of the empirical exercises using the factor model. For each exercise, the data vintages used, the specification of the forecasting model with and without sentiment, as well as the estimation and evaluation sample periods are described.
[a]The last observation in the dataset contains actual values for all the variables.
[b]Depending on the date for which a real-time dataset is created, the appropriate vintage is used. For example, the dataset representing the information available to forecasters in March 2005 uses data vintages from March 2005.
[c]Sentiment variable is removed completely from the model.
[d]This is the sample for the first of a series of recursive regressions.
[e]Date range represents the evaluation of nowcasts (horizon 0). The evaluation sample for longer horizon forecasts change according to horizon. For example, evaluation sample of forecasts with horizon 1 starts one month/quarter later.

DOI: 10.1002/jae

3.3. Marginal Impact of Consumer Confidence on Forecast Accuracy

Using in-sample fitted values, we first measure the difference in mean squared errors (MSEs) between models of consumption expenditure with and without the entire series of a confidence measure. This is the first exercise discussed above. We do so for all possible pairs of confidence measures and types of consumption expenditure. The results are presented in Table III, where a relative MSE value that is smaller than 1 means that forecasts using consumer confidence have a smaller MSE. These cells are shaded. In this and all subsequent exercises, we test the reported differences in MSEs using the Diebold and Mariano (1995) test with its small-sample modification by Harvey *et al.* (1997).[15] Whenever the difference between two competing MSEs with and without using a consumer confidence measure is statistically significant at 10%, we report the relative MSE in bold. We also report the RMSE of the benchmark model, i.e. factor model without confidence measure, for both the monthly and quarterly models.

We observe that adding a confidence measure improves the accuracy of the forecasts consistently, and for services and total consumption this improvement is statistically significant. On average, adding a consumer confidence measure reduces the in-sample MSE by about 8%. The models of services and total consumption benefit the most from this addition, with reductions in MSE at 15% and 10% respectively. We also find that the improvements due to the addition of a University of Michigan sentiment measure (ICS) are similar to those obtained using its Conference Board counterpart (CCI). While improvements in forecast accuracy are noted for all types of consumption, improvements in consumption of durable and non-durable goods are generally insignificant, e.g. p-values on average are around 0.1 in the quarterly models. This is in contrast to our findings regarding consumption of services. Quite clearly, the message from the previous subsection is not only confirmed but also reinforced here: even when considered in a rich information context, confidence matters when forecasting consumption.

We then proceed with the second exercise, where we forecast consumption expenditure using a series of reconstructed datasets (with the first one ending in January 1995) that reflect the varying data release schedules and publication lags across explanatory variables. Each month, for each pair of consumption and confidence, we make two sets of forecasts: one before we observe the confidence measure for that very month, and one after we observe it. By comparing the MSEs of the two sets of forecasts, we reveal the value of release timing of the latest measure of confidence in forecasting consumption.

With quarterly data, we place ourselves in six different points in time every quarter, namely the day of each of the 3 months of the quarter when the confidence measure is released (last Tuesday of the month for the Conference Board confidence measure and last Friday of the month for the University of Michigan measure). On each of the three Tuesdays/Fridays, we make 10 forecasts. Five of them (horizons 0–4) are based on the information set that includes this latest release of confidence measure, and the other five are based on the information set that does not. With monthly data, we make 14 forecasts at the end of each month, seven of them (horizon 0–6) with the latest release of the confidence measure and the other seven without. For both quarterly and monthly models, horizon 0 always refers to the current period (quarter/month), horizon 1 refers to the immediate next quarter/month, and so on. Our estimation samples start from January 1982 and our evaluation samples cover January 1995 to June 2014.

The results from this exercise are summarized in Table IV, where relative MSEs are reported for six monthly horizons with monthly data and up to four quarters with quarterly data. We observe that for both quarterly and monthly models, in 56% of the cases, the forecasts made with the latest confidence

[15] The standard Diebold–Mariano test gives conservative results in MSE comparisons between nested models. Therefore, the true contribution of confidence measures could be more significant than that indicated by the p-values reported in Tables 3 and 4 (see Clark and McCracken 2011). In addition, one may be concerned about the fact that our forecasts are based on generated regressors, i.e. the estimated factors. However, as Banerjee *et al.* (2008) show, forecasts from true and estimated factors are in general very similar in the presence of structural breaks with N and T that are even smaller than ours.

DOI: 10.1002/jae

Table III. Relative MSEs of in-sample predictions with/without consumer sentiment/confidence

PCE and component	Data series	Sentiment	Monthly model: Benchmark RMSE	Monthly model: Relative MSE	Monthly model: p-value	Quarterly model (revised data): Benchmark RMSE	Quarterly model (revised data): Relative MSE	Quarterly model (revised data): p-value
Durable goods	Conference Board	Expectations	31.206	0.999	0.827	9.386	0.958	0.245
		Overall index		0.998	0.718		0.975	0.340
	University of Michigan	Expectations		0.998	0.620		0.961	0.229
		Overall index		0.998	0.736		0.969	0.298
Non-durable goods	Conference Board	Expectations	7.547	0.993	0.806	2.265	0.946	0.200
		Overall index		0.991	0.659		0.929	0.185
	University of Michigan	Expectations		0.991	0.650		0.919	0.131
		Overall index		0.991	0.613		0.910	0.106
Services	Conference Board	Expectations	3.256	**0.935**	**0.009**	1.427	**0.732**	**0.000**
		Overall index		**0.958**	**0.044**		**0.819**	**0.012**
	University of Michigan	Expectations		**0.936**	**0.011**		**0.755**	**0.002**
		Overall index		**0.933**	**0.007**		**0.741**	**0.001**
Total	Conference Board	Expectations	5.592	0.987	0.162	1.897	**0.804**	**0.005**
		Overall index		0.987	0.168		**0.858**	**0.022**
	University of Michigan	Expectations		**0.981**	**0.086**		**0.803**	**0.004**
		Overall index		0.982	0.107		**0.805**	**0.005**

Note: This table shows the relative MSE of in-sample predictions made with and without consumer sentiment/confidence measure. Relative MSE is the ratio between models with sentiment and models without sentiment; i.e. relative MSE smaller than 1 means adding sentiment measure improves the fit/predictive performance. Shaded cells are those in which the relative MSE is smaller than 1. The relative MSEs and p-values in bold are those where *one-sided* DM test rejects at 10%. Benchmark RMSE is that from the benchmark model where no confidence measure is included. The sample covers January 1982 to June 2014.

measure have lower MSEs (with the average improvement being about 1.2% in monthly models and 3.5% in quarterly models). In most cases where an improvement in out-of-sample forecasting performance is observed, the dependent variable is either services or total consumption, especially when this improvement is statistically significant. This is consistent with the results from the previous exercise. While forecasts of durable and non-durable goods consumption also benefit from the inclusion of the confidence measures, the effects are smaller. It is perhaps not surprising that consumption of services, which is discretionary in nature, and has evolved as a major component (over two-thirds) of aggregate consumption, is better predicted by consumer sentiment.

It should be mentioned here that the improvement we get when we include the latest value of confidence is often not statistically significant. However, let us recall that in this exercise the forecast improvements are based on an information set that is augmented in a very marginal way, i.e. with just one extra observation. Overall, even with this qualification, the broad picture that emerges from the results of this and the previous subsection is still one where confidence measures often lead to noticeable improvements in the accuracy of consumption forecasts. Such improvements can be attributed to both the information content of confidence measures and the timeliness of their releases. However, the effect of data revisions, i.e. the third determinant as discussed before, remains unaccounted for.

3.4. Forecasting Consumption in Real Time

In this exercise, we consider the most realistic set-up, which allows us to examine the effect of all three determinants affecting the role of consumer confidence in forecasting consumption expenditure, i.e. the information content, the timing and lag of data releases, and data revisions. So, in contrast to the pseudo-real-time exercises of the previous subsection, here we create a series of true real-time datasets from March 2005 to June 2014 for the last Friday of each month for the University of Michigan sentiment measures, and for the last Tuesday of each month for the Conference Board confidence measures. As discussed earlier, we achieve this using the stylized data release schedule and publication lags, as well as information on data revisions associated with the GRS dataset. We repeat this exercise for both quarterly and monthly consumption data, just as with all previous exercises.

On each of the 113 Fridays (or Tuesdays) in our sample, we produce 10 quarterly forecasts and 14 monthly forecasts, similar to the second exercise in the previous subsection, with the estimation sample starting from January 1982. Table V presents the results based on evaluation samples from March 2005 to June 2014. In a clear majority of cases, adding a confidence measure notably improves the real-time forecasting performance (62% of the time for quarterly models, with an average improvement of 2.8% and 51% of the time for monthly models, with an average improvement of 2.0%). In 16% of all cases, this improvement is statistically significant.

Of course, the same observation applies here as with the previous exercise. We find it remarkable that augmenting such a large information set in such a marginal way (i.e. the innovation implied by the latest announcement over the previous one) often leads to noticeable improvements in the forecasts. The obvious next step is to examine, in this realistic context as well, the effect of the entire confidence variable on the forecasts. Table VI reports the results from an exercise (otherwise identical to that reported in Table V) where forecasts based on information sets that include a confidence variable are compared to those based on information sets without any confidence measure, so not just the latest observation but the entire series of confidence measure is removed. Thus this exercise considers the value of the complete time series of the confidence measure against the scenario where consumer sentiment never existed at all. The evidence is overwhelming: For 90% of the time for quarterly models, adding a confidence measure leads to improvements (with an average improvement over all types of consumption of about 25%). Similarly, for monthly models, adding a confidence measure leads to improvements for 88% of the time (with an average improvement of 9%). In 65% of all cases, this improvement is statistically significant, while deteriorations that are statistically significant are very

DOI: 10.1002/jae

Table IV. Relative MSEs of out-of-sample predictions with/without the latest consumer sentiment/confidence

PCE and component	Data series	Sentiment	Monthly							Quarterly														
			H = 0	H = 1	H = 2	H = 3	H = 4	H = 5	H = 6	First month of a quarter					Second month of a quarter					Third month of a quarter				
										H = 0	H = 1	H = 2	H = 3	H = 4	H = 0	H = 1	H = 2	H = 3	H = 4	H = 0	H = 1	H = 2	H = 3	H = 4
Durable goods	Conference	Expectations	**0.989**	**0.977**	1.023	0.996	**1.004**	1.000	**1.005**	1.064	0.985	0.992	1.039	**0.962**	1.060	0.987	0.995	1.035	**0.959**	1.081	0.986	1.002	1.030	**0.954**
	Board	Overall index	**0.987**	1.002	**0.985**	1.015	**0.990**	**1.023**	**0.976**	0.954	1.048	1.007	**1.151**	1.046	0.962	1.043	1.009	**1.149**	1.046	0.964	1.040	1.001	**1.147**	1.045
	University of	Expectations	**0.972**	1.003	**1.014**	0.998	**0.990**	**1.009**	**0.992**	**0.945**	1.035	0.944	**1.094**	1.014	0.962	1.036	0.947	**1.095**	1.014	0.980	1.042	0.946	**1.096**	1.014
	Michigan	Overall index	0.977	1.003	1.007	1.004	**0.995**	1.000	1.005	0.994	**0.970**	0.951	1.050	1.015	1.006	**0.972**	0.957	1.049	1.015	1.031	0.978	0.960	1.048	1.015
Non-durable goods	Conference	Expectations	0.988	1.001	1.005	0.999	1.004	1.004	**1.004**	0.959	0.982	1.009	0.985	1.035	0.962	0.984	1.010	0.986	1.026	0.985	0.990	1.015	0.989	1.017
	Board	Overall index	0.994	1.003	0.993	1.007	**0.989**	**1.010**	1.001	0.970	1.025	0.994	1.005	0.930	0.974	1.024	0.998	1.006	0.931	0.979	1.027	1.001	1.006	0.932
	University of	Expectations	0.992	1.000	1.000	0.998	0.998	**1.005**	0.997	**0.948**	**0.978**	**1.051**	0.994	**0.929**	**0.956**	0.982	**1.051**	0.995	**0.931**	0.969	0.984	**1.053**	0.998	**0.932**
	Michigan	Overall index	0.992	1.002	0.999	0.999	1.004	1.001	1.001	**0.950**	**0.974**	**1.032**	**0.975**	0.955	**0.958**	0.977	**1.033**	0.976	0.954	0.973	0.980	**1.036**	0.979	0.953
Services	Conference	Expectations	**0.970**	0.998	1.018	1.000	**0.989**	1.002	1.005	0.987	0.990	0.966	**0.946**	**0.920**	0.992	1.000	0.970	**0.950**	**0.921**	1.004	1.009	0.980	**0.956**	**0.924**
	Board	Overall index	0.986	1.000	1.017	**1.007**	0.996	1.002	1.000	1.024	1.053	0.994	0.974	0.953	1.026	1.057	0.998	0.975	0.955	**1.034**	1.058	1.002	0.977	0.956
	University of	Expectations	0.990	1.006	1.017	1.001	0.984	1.001	0.999	**1.026**	1.001	**0.968**	**0.916**	0.994	1.032	1.008	**0.972**	0.917	0.998	**1.040**	1.014	**0.976**	0.922	1.000
	Michigan	Overall index	0.987	0.992	1.015	1.003	0.997	**0.988**	1.004	1.006	1.005	**0.976**	**0.918**	0.989	1.010	1.012	**0.979**	**0.919**	0.991	1.016	1.016	0.983	**0.924**	0.992
Total	Conference	Expectations	**0.968**	**0.976**	**1.045**	**0.984**	1.003	1.002	1.013	0.961	0.952	0.963	1.011	0.944	0.968	0.961	0.971	1.008	0.941	1.007	0.971	0.983	1.008	0.940
	Board	Overall index	0.994	0.996	0.987	**1.021**	**0.980**	**1.029**	**0.979**	0.956	1.085	1.037	**1.079**	0.965	0.962	1.087	1.042	**1.079**	0.969	0.966	1.086	1.043	**1.078**	0.970
	University of	Expectations	**0.974**	1.004	**1.015**	0.994	**0.978**	1.017	**0.988**	**0.895**	1.022	0.965	1.037	0.967	**0.907**	1.029	0.970	1.039	0.970	0.921	1.036	0.973	1.044	0.970
	Michigan	Overall index	**0.976**	**0.993**	**1.014**	0.997	0.995	0.999	**1.006**	**0.926**	**0.960**	0.962	0.991	0.992	**0.938**	0.967	0.969	0.993	0.995	0.961	0.974	0.975	0.997	0.996

Note: This table shows the relative MSE of out-of-sample predictions made with and without the latest release (i.e. one value) of consumer sentiment/confidence measure. Relative MSE is the ratio between models with sentiment and models without sentiment; i.e. relative MSE smaller than 1 means adding sentiment measure improves the fit/predictive performance. Shaded cells are those in which the relative MSE is smaller than 1. The p-values in bold mean two-sided DM test rejection at 10%. Training sample starts from January 1982. Evaluation sample covers January 1995 to June 2014.

Table V. Relative MSEs of real-time forecasts with/without the latest consumer sentiment/confidence

			Monthly							Quarterly														
			H = 0	H = 1	H = 2	H = 3	H = 4	H = 5	H = 6	First month of a quarter					Second month of a quarter					Third month of a quarter				
PCE and component	Data series	Sentiment								H = 0	H = 1	H = 2	H = 3	H = 4	H = 0	H = 1	H = 2	H = 3	H = 4	H = 0	H = 1	H = 2	H = 3	H = 4
Durable goods	Conference	Expectations	**0.972**	0.993	1.021	0.982	0.995	1.013	**1.025**	0.938	0.955	**0.949**	1.019	1.018	1.031	**0.961**	0.996	1.003	0.995	1.047	**0.956**	0.996	1.013	0.995
	Board	Overall index	0.988	1.001	1.005	0.994	0.999	1.000	1.003	0.976	0.984	0.986	1.001	1.003	1.000	**0.988**	0.999	1.001	1.002	1.005	**0.985**	0.999	1.002	0.998
	University of	Expectations	**0.985**	**0.987**	1.011	1.002	**0.986**	1.013	0.999	0.971	0.988	0.973	0.995	1.024	1.028	0.981	0.983	0.987	1.009	1.026	**0.967**	0.991	0.997	1.013
	Michigan	Overall index	**0.979**	0.998	1.016	0.996	**0.984**	1.017	0.996	0.967	1.001	0.983	1.007	1.036	1.014	0.983	0.985	0.993	1.016	**1.031**	**0.970**	0.989	1.000	1.013
Non-durable goods	Conference	Expectations	**0.965**	**1.055**	0.994	1.010	**1.018**	**0.980**	1.024	0.937	1.036	**0.930**	1.001	**1.011**	1.001	1.033	0.958	1.024	1.002	0.995	1.045	0.972	1.026	0.986
	Board	Overall index	**0.980**	**1.027**	0.994	1.004	1.004	**0.984**	1.003	0.967	1.007	0.972	0.991	**1.001**	1.003	1.007	0.979	0.997	**1.000**	0.996	1.009	0.983	0.995	**0.992**
	University of	Expectations	**0.983**	**1.027**	**0.979**	**1.014**	0.994	1.002	**1.015**	0.978	0.985	0.998	0.981	**0.999**	1.013	1.003	1.002	0.999	0.979	1.004	0.998	0.995	1.003	0.982
	Michigan	Overall index	**0.987**	**1.034**	**0.965**	**1.033**	**0.976**	1.013	1.004	0.973	0.998	0.992	0.999	**1.006**	1.002	1.003	0.994	0.996	**0.972**	1.004	0.999	0.990	0.998	0.972
Services	Conference	Expectations	1.047	1.002	**1.035**	1.020	0.954	1.003	0.994	1.137	1.087	**0.811**	0.923	1.170	1.208	1.058	**0.924**	1.001	1.078	**1.124**	1.019	**0.914**	0.979	**1.085**
	Board	Overall index	**0.977**	**0.982**	**1.009**	0.994	**0.979**	0.992	0.993	0.993	1.001	**0.930**	0.976	1.025	1.033	1.012	**0.974**	0.997	1.009	1.015	1.005	**0.969**	0.990	1.010
	University of	Expectations	1.005	0.979	1.011	1.006	0.975	0.986	0.993	1.029	0.895	0.957	0.998	**0.924**	**1.099**	0.954	1.022	0.994	**0.961**	1.046	0.934	0.985	0.971	**0.964**
	Michigan	Overall index	1.008	**0.969**	**1.039**	0.980	0.978	1.008	**0.970**	1.077	**0.887**	0.940	1.014	**0.941**	**1.104**	0.962	0.983	0.989	**0.959**	1.039	**0.934**	0.983	0.977	**0.962**
Total	Conference	Expectations	**0.915**	1.016	1.033	0.972	1.002	1.014	1.040	0.910	0.972	**0.853**	0.996	1.070	1.127	0.981	0.961	1.013	1.012	**1.121**	0.974	0.970	1.012	1.011
	Board	Overall index	**0.951**	**1.013**	1.010	0.987	0.997	**0.991**	1.009	0.951	0.985	**0.946**	0.995	1.003	1.012	0.993	0.979	0.999	0.996	1.010	0.988	0.980	0.995	0.991
	University of	Expectations	**0.956**	**0.980**	1.004	**1.019**	**0.967**	1.008	1.005	0.954	0.953	0.946	1.001	0.976	1.086	0.973	0.977	0.997	0.964	1.066	0.946	0.973	0.996	0.972
	Michigan	Overall index	**0.954**	**0.990**	**1.017**	1.007	**0.957**	**1.029**	0.980	0.962	0.979	0.947	1.030	0.995	1.062	0.975	0.962	0.997	0.958	1.061	**0.947**	0.969	0.998	0.963

Note: This table shows the relative MSE of real-time forecasts made with and without the latest release (i.e. one value) consumer sentiment/confidence measure. Relative MSE is the ratio between models with sentiment and models without sentiment; i.e. relative MSE smaller than 1 means adding sentiment measure improves the fit/predictive performance. Shaded cells are those in which the relative MSE is smaller than 1. The p-values in bold mean two-sided DM test rejection at 10%. Training sample starts from January 1982. Evaluation sample covers March 2005 to June 2014.

rare: only 1.4% of the time. In this scenario, in addition to services, non-durable and total consumption are also very significant.

To further understand the contribution of consumer confidence measures to the accuracy of consumption forecasts, and in particular possibly changing patterns of such contributions over the business cycle, we proceed by separately evaluating the real-time forecasts over the recessionary period 2007:12–2009:6, the pre-recessionary period of 2005:3–2007:11, and the post-recessionary period 2009:07–2014:6. For each period, we report the relative MSE of the forecasts made with and without confidence measures (see Table VII). Like the exercise above, we consider both the marginal contribution of the latest observation of confidence measures (top panel) and the overall contribution of the confidence variable (bottom panel). For brevity, for quarterly forecasts, we only report the results obtained using forecasts made in the third month of a quarter.

We observe that even when evaluated over recession and non-recession periods separately, forecasts made with consumer confidence information are very often more accurate, and the effects are quite discernible with monthly data. Note that with relatively small subsamples the use of monthly data helps to estimate the effects more precisely than those obtainable with quarterly observations. We also observe that the contribution of confidence measures is systematically higher during the recession compared to the non-recession periods, and are generally significant for all categories of consumption including durables. As before, forecasts for services consumption benefit the most. However, the additional sensitivity of aggregate consumption to sentiment during the last recession is partly attributable to its sizable effect on durables (including autos), which is known to be the most volatile component of total consumption. Using a VAR regime-switching model, Ivanova and Lahiri (2001) show that the benefit of including sentiment is larger in periods like (but not exclusively) recessions when conflicting economic and social-political news cause high overall uncertainty and wide swings in near-term expectations in personal income, and hence wide changes in discretionary consumer spending on durables. We now find a similar effect via services.

At this point, we should underscore one of our novel empirical findings that is repeated in both this and previous exercises. Looking across components of personal consumption expenditures, the sizable improvements to the forecasting performance in total consumption come often through services consumption. Over 1950:01–2014:10, whereas the share of durable consumption has hovered around 13%, the proportion of non-durable consumption in total consumption has steadily fallen from 45% to 22%, and the same for services has increased from 39% to 67%. The latter component includes housing and utilities, health care, social assistance, education, transportation, recreation, food and accommodation, information, finance and insurance, and the like. Health care and a few other service items like transportation, recreation, finance, and insurance are also known to be highly procyclical. These aspects of services consumption may explain the effect of sentiment over phases of the business cycle.

Furthermore, and in light of the evidence of temporal instabilities being a prevalent concern in macroeconomic forecasting (see Rossi, 2013), we investigate the robustness of our results also with respect to temporal changes of different sorts from those suggested by the above subsample analysis. In particular, we consider a series of estimation and forecasting exercises using rolling regressions (with window sizes of 10, 15, and 20 years); the forecasting results (not reported here for brevity) were very similar to those in Table VI, where services came out to be the most dominant component. A time-varying (random walk) specification for the coefficient relating confidence to total consumption indicated a very slow but smooth decline of the coefficient from around 0.079 to 0.069. After all, not only has the relative importance of services as a component of aggregate consumption been changing over time, but its own composition has also changed over the sample. Strikingly, spending on health care now accounts for over 25% of spending on services. In order to study the time variation in the relative predictive performance of the sentiment variables over the entire sample, and test its stability over time, we also employ the Giacomini and Rossi (2010) fluctuation test. Figure 2 presents the time

Table VI. Relative MSEs of real-time forecasts with/without the consumer sentiment/confidence variable

			Monthly							Quarterly														
										First month of a quarter					Second month of a quarter					Third month of a quarter				
PCE and component	Data series	Sentiment	H = 0	H = 1	H = 2	H = 3	H = 4	H = 5	H = 6	H = 0	H = 1	H = 2	H = 3	H = 4	H = 0	H = 1	H = 2	H = 3	H = 4	H = 0	H = 1	H = 2	H = 3	H = 4
Durable goods	Conference Board	Expectations	0.972	0.998	1.004	0.984	1.002	**1.007**	0.990	0.820	0.937	0.900	0.999	1.140	0.927	0.932	0.932	0.969	1.077	1.018	0.919	0.949	0.962	1.042
		Overall index	0.990	1.003	1.000	0.997	1.000	1.000	0.995	0.977	1.030	1.002	0.998	1.043	1.009	1.042	1.016	0.996	1.028	1.027	1.045	1.023	0.996	1.017
	University of Michigan	Expectations	0.980	0.996	1.007	0.998	0.994	**1.007**	0.991	0.902	0.965	0.949	0.991	1.069	0.960	0.983	0.956	0.979	1.043	1.014	0.965	0.964	0.973	1.016
		Overall index	0.985	1.012	1.012	0.999	0.999	**1.014**	0.991	0.901	0.986	0.968	1.002	1.083	0.949	1.003	0.973	0.990	1.051	1.011	0.986	0.978	0.981	1.015
Non-durable goods	Conference Board	Expectations	0.982	**1.027**	**0.973**	**0.978**	**0.966**	**0.950**	0.977	**0.834**	**0.868**	**0.786**	**0.828**	**0.872**	0.895	0.887	0.804	**0.844**	**0.862**	0.918	0.896	0.818	**0.852**	**0.847**
		Overall index	**0.965**	**0.987**	**0.957**	**0.959**	**0.957**	**0.953**	**0.971**	**0.865**	0.897	**0.891**	**0.919**	0.952	0.903	0.913	0.904	**0.934**	0.953	0.905	0.927	0.902	**0.933**	**0.947**
	University of Michigan	Expectations	**0.967**	**0.989**	**0.965**	**0.982**	**0.976**	**0.980**	**0.982**	0.855	0.847	0.882	**0.826**	**0.871**	0.887	0.870	0.893	**0.847**	**0.863**	0.900	0.876	0.887	**0.844**	**0.855**
		Overall index	**0.967**	**0.981**	**0.951**	**0.979**	**0.957**	**0.977**	**0.970**	0.819	0.831	**0.849**	**0.828**	**0.881**	0.846	0.851	0.860	**0.842**	**0.862**	0.865	0.860	**0.856**	**0.832**	**0.844**
Services	Conference Board	Expectations	**0.871**	**0.828**	**0.817**	**0.791**	**0.771**	**0.811**	**0.801**	**0.524**	**0.401**	**0.414**	**0.503**	**0.507**	**0.610**	**0.420**	**0.426**	**0.533**	**0.517**	**0.632**	**0.456**	**0.409**	**0.534**	**0.536**
		Overall index	**0.848**	**0.849**	**0.862**	**0.843**	**0.841**	**0.864**	**0.860**	**0.610**	**0.533**	**0.595**	**0.637**	**0.644**	**0.673**	**0.548**	**0.589**	**0.651**	**0.650**	**0.725**	**0.588**	**0.560**	**0.655**	**0.660**
	University of Michigan	Expectations	**0.816**	**0.804**	**0.825**	**0.810**	**0.801**	**0.825**	**0.830**	**0.518**	**0.435**	**0.516**	**0.570**	**0.519**	**0.569**	**0.450**	**0.519**	**0.570**	**0.538**	**0.572**	**0.467**	**0.494**	**0.552**	**0.544**
		Overall index	**0.817**	**0.798**	**0.828**	**0.790**	**0.800**	**0.823**	**0.808**	**0.517**	**0.399**	**0.468**	**0.533**	**0.479**	**0.572**	**0.420**	**0.458**	**0.530**	**0.494**	**0.589**	**0.452**	**0.448**	**0.521**	**0.502**
Total	Conference Board	Expectations	**0.839**	**0.914**	**0.903**	**0.871**	**0.888**	**0.886**	**0.849**	**0.528**	**0.614**	**0.560**	**0.665**	0.698	**0.667**	**0.632**	**0.615**	**0.665**	**0.675**	**0.745**	**0.646**	**0.639**	**0.666**	**0.661**
		Overall index	**0.889**	**0.941**	**0.927**	**0.909**	**0.913**	**0.914**	**0.897**	**0.740**	**0.792**	**0.768**	**0.797**	0.810	**0.796**	**0.815**	**0.799**	**0.806**	0.808	**0.816**	**0.841**	**0.804**	**0.812**	**0.803**
	University of Michigan	Expectations	**0.852**	**0.896**	**0.918**	**0.909**	**0.887**	**0.916**	**0.888**	**0.605**	**0.668**	**0.668**	**0.693**	**0.684**	**0.680**	**0.690**	**0.692**	**0.696**	**0.683**	**0.698**	**0.685**	**0.695**	**0.692**	**0.675**
		Overall index	**0.860**	**0.907**	**0.921**	**0.896**	**0.882**	**0.919**	**0.866**	**0.590**	**0.677**	**0.650**	**0.692**	**0.690**	**0.656**	**0.697**	**0.668**	**0.689**	**0.677**	**0.689**	**0.699**	**0.679**	**0.686**	**0.659**

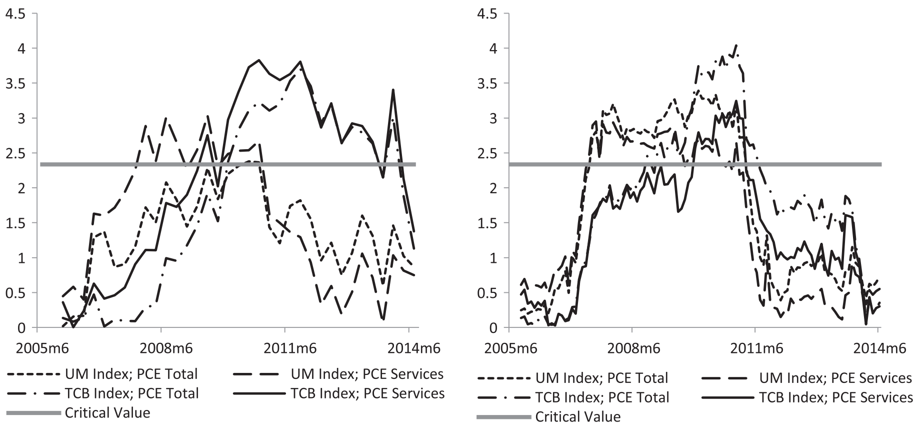
Note: This table shows the relative MSE of real-time forecasts made with and without the consumer sentiment/confidence measure (i.e. the sentiment variable, not just the last observation of the variable). Relative MSE is the ratio between models with sentiment and models without sentiment; i.e. relative MSE smaller than 1 means adding sentiment measure improves the fit/predictive performance. Shaded cells are those in which the relative MSE is smaller than 1. The p-values in bold mean one-sided DM test rejection at 10%. Training sample starts from January 1982. Evaluation sample covers March 2005 to June 2014.

Table VII. Evaluation of real-time forecasts over recession and non-recession periods

Sentiment	Horizon	Durable Goods Consumption						Non-Durable Goods Consumption						Services Consumption						Total Consumption					
		Monthly			Quarterly			Monthly			Quarterly			Monthly			Quarterly			Monthly			Quarterly		
		05:3-07:11	07:12-09:6	09:7-14:6	05:3-07:11	07:12-09:6	09:7-14:6	05:3-07:11	07:12-09:6	09:7-14:6	05:3-07:11	07:12-09:6	09:7-14:6	05:3-07:11	07:12-09:6	09:7-14:6	05:3-07:11	07:12-09:6	09:7-14:6	05:3-07:11	07:12-09:6	09:7-14:6	05:3-07:11	07:12-09:6	09:7-14:6
With/without the latest observation																									
TCB Index	0	0.996	**0.838**	1.012	1.011	1.009	0.994	0.997	**0.960**	0.987	0.982	1.001	0.995	1.003	**0.827**	0.988	0.998	1.029	1.010	0.988	**0.842**	0.981	1.010	1.019	0.993
	1	1.000	0.997	1.004	0.977	0.989	0.985	**0.994**	**1.063**	1.016	0.998	1.013	1.006	0.995	**0.888**	0.995	1.006	0.976	1.015	**0.992**	1.019	**1.032**	0.949	0.990	0.994
	2	**0.996**	1.102	0.987	**1.004**	0.965	1.023	**1.000**	**0.982**	1.003	**1.030**	0.971	0.983	1.006	1.000	1.014	0.987	0.964	0.970	**0.999**	**1.040**	0.997	**1.063**	0.955	0.990
	3	1.004	**0.911**	1.009	0.999	1.006	0.997	**1.007**	1.015	0.990	**0.994**	**0.987**	1.012	1.000	0.931	1.005	1.105	**0.982**	0.990	1.015	**0.947**	0.991	0.988	**0.989**	1.006
	4	1.000	0.982	1.004	**1.005**	0.993	1.003	1.012	0.997	1.006	**0.986**	0.988	1.006	**0.984**	0.918	0.990	0.975	0.972	1.026	1.000	**0.967**	1.021	**0.980**	0.984	**1.007**
	5	**1.000**	1.008	0.999				0.991	**0.976**	0.989				**1.019**	**0.964**	0.978				1.002	**0.975**	0.992			
	6	0.997	1.015	1.003				0.996	1.005	1.007				0.999	**0.957**	0.996				0.991	1.007	1.023			
UM Index	0	0.990	**0.795**	1.024	1.041	1.028	1.024	0.995	0.986	**0.981**	0.947	1.030	0.984	0.994	0.885	1.040	0.963	1.082	1.027	**0.953**	**0.847**	1.041	1.039	1.086	1.031
	1	**1.009**	**0.961**	0.998	0.935	0.972	0.984	0.988	1.079	**1.022**	1.004	**0.994**	1.007	0.988	0.806	0.985	0.899	0.954	0.931	**0.991**	0.969	1.006	0.851	0.969	0.953
	2	0.988	**1.185**	0.998	**1.028**	0.896	1.035	1.010	**0.920**	0.986	**1.059**	**0.989**	**0.955**	1.022	**1.232**	1.020	0.886	0.927	0.998	1.002	**1.072**	0.990	**1.198**	**0.901**	0.987
	3	1.002	**0.972**	0.997	0.972	1.033	0.976	1.009	**1.067**	1.012	**1.011**	0.996	0.997	0.969	0.863	1.010	**1.361**	1.019	0.961	1.018	0.994	1.005	0.944	1.019	0.982
	4	0.995	**0.879**	1.006	1.043	0.953	1.080	1.002	**0.953**	**0.984**	**0.949**	0.954	1.039	**0.971**	**0.823**	1.016	**0.833**	0.885	**0.993**	0.999	**0.870**	0.987	0.958	0.924	1.047
	5	1.003	**1.139**	0.998				1.001	**1.035**	0.998				**1.041**	0.947	0.993				1.002	**1.073**	1.025			
	6	0.991	0.973	1.005				0.997	1.001	1.012				1.008	**0.867**	0.962				0.983	**0.954**	0.998			
With/without the entire history of the variable																									
TCB Index	0	0.996	**0.790**	1.029	1.003	0.886	**1.343**	1.002	**0.938**	0.963	1.027	0.911	0.744	1.006	**0.580**	**0.813**	1.012	1.078	**0.621**	0.995	**0.719**	**0.891**	1.030	0.872	**0.626**
	1	1.003	**0.978**	**1.012**	0.993	0.926	1.193	**1.005**	**0.975**	0.988	1.060	0.920	0.869	1.006	**0.657**	**0.802**	1.029	0.947	**0.501**	**1.008**	**0.872**	**0.925**	1.055	0.882	0.759
	2	1.000	**0.993**	1.003	1.001	**0.932**	**1.110**	1.006	**0.915**	**0.968**	1.031	0.847	0.963	**1.005**	**0.745**	**0.806**	1.058	0.693	**0.531**	**1.013**	**0.871**	**0.882**	1.021	0.810	**0.770**
	3	1.004	**0.897**	1.021	0.997	**0.987**	1.010	**1.010**	**0.924**	**0.962**	0.994	0.921	**0.934**	0.999	**0.701**	**0.784**	**0.942**	0.749	**0.628**	**1.015**	**0.813**	**0.885**	**0.985**	0.893	**0.693**
	4	1.000	**0.962**	1.013	1.012	0.953	1.127	1.003	**0.924**	**0.964**	1.011	0.980	**0.845**	**0.999**	**0.724**	**0.773**	**0.955**	0.819	**0.610**	1.001	**0.851**	**0.878**	**1.029**	0.928	0.611
	5	0.999	**0.967**	1.011				0.990	**0.923**	0.962				**1.014**	**0.801**	**0.787**				1.000	**0.874**	**0.865**			
	6	**0.997**	**0.950**	1.009				0.998	0.949	0.977				**0.993**	**0.813**	**0.791**				0.990	**0.886**	**0.855**			
UM Index	0	**0.991**	**0.795**	1.037	0.906	**0.738**	**1.730**	0.998	0.951	0.955	0.986	0.864	0.729	0.966	**0.445**	**0.814**	**0.576**	0.894	**0.514**	**0.961**	**0.971**	0.900	0.718	0.714	0.635
	1	1.003	**1.061**	1.007	**0.989**	0.818	1.158	1.002	**0.964**	0.984	**1.158**	0.831	0.769	0.991	**0.455**	**0.772**	**0.703**	0.766	**0.384**	1.006	**0.810**	**0.882**	**1.047**	0.709	0.624
	2	0.989	**1.125**	1.003	0.998	**0.763**	1.153	**1.011**	**0.906**	**0.959**	**1.123**	**0.806**	0.843	0.999	**0.579**	**0.790**	1.002	**0.453**	**0.436**	**1.012**	**0.868**	**0.870**	**1.106**	**0.621**	0.670
	3	1.002	**0.944**	1.013	1.002	**0.922**	1.056	**1.005**	0.967	**0.974**	0.968	0.805	**0.836**	**0.971**	**0.437**	**0.770**	0.737	0.514	**0.519**	1.013	**0.773**	**0.884**	0.978	0.736	**0.586**
	4	0.999	**0.940**	1.017	**0.990**	0.830	1.334	0.999	**0.932**	0.957	**1.015**	0.861	0.745	1.003	**0.479**	**0.756**	0.658	0.549	**0.486**	0.996	**0.770**	**0.864**	**0.966**	0.734	0.527
	5	1.003	**1.046**	**1.016**				0.998	0.967	0.975				**1.032**	**0.597**	**0.750**				0.997	**0.872**	**0.885**			
	6	0.997	**0.904**	1.016				0.996	**0.944**	0.983				0.989	**0.617**	**0.742**				**0.983**	**0.802**	**0.853**			

Note: This table reports the relative MSEs of forecasts made in real time using real-time data with and without sentiment information (i.e. the third exercise). The forecasts are evaluated over the 2005:3 to 2007:11 non-recession period, the 2007:12 to 2009:6 recession periods, and then the 2009:7 to 2014:6 non-recession periods. A relative MSE value smaller than 1 means forecasts made with sentiment are more accurate. Monthly forecasts are made at the end of the month. Quarterly forecasts are made at the end of the 3rd month of a quarter. Cells with relative MSE smaller than 1 are shaded. The p-values in bold mean DM test rejection at 10%. The test is two-sided for the upper panel of the table and one-sided for the lower panel.

DOI: 10.1002/jae



(a) Quarterly Model　　**(b)** Monthly Model

Figure 2. Fluctuation test statistic and critical value. This figure shows the fluctuation test statistic calculated using real-time current-period forecasts made with/without sentiment variable based on rolling window regressions with a 20-year rolling window. The left plot shows the results from quarterly models and the right plot shows the results from monthly models

profile of the test statistic calculated using real-time (current period) forecasts based on 20-year rolling window regressions. Figure 2(a) shows the results from quarterly models and Figure 2(b) shows the same for monthly models for services and total consumption using both TCB and Michigan indices. The 10% one-sided critical value is specified by the solid horizontal line. Theoretically, the test reveals whether at each point in time the sentiment indices have zero additional forecasting power against the alternative that sentiment contributes at least at some point over the sample. In our context, the advantage of the test is that we do not need to specify the nature of the instability under the alternative hypothesis. The interesting finding here is that the lines do not cross the critical values exactly over the latest recessionary period 2007–2009. Many of the values exceed the critical value much before the cyclical peak of December 2007 and remained significant long after the cyclical trough of June 2009. The results do suggest time variation in the predictive power of sentiment. Consistent with the results of our subsample analysis discussed above, both the Michigan and TCB indices are most useful around recession periods, in our case, from 2006 to 2011.

3.5. Discussion

Overall, the main results from these exercises establish a positive effect of consumer confidence on total consumption, which represents a departure from the PI/RE hypothesis. Past literature has considered two hypotheses to explain such excess sensitivity of consumption to sentiment: (i) precautionary savings motive; and (ii) that sentiment captures household expectations of income growth.

However, due to many methodological reasons, the use of time series data has not provided strong evidence in favor of the first hypothesis. Souleles (2004), using matched household-level data, finds substantial evidence of a precautionary savings channel, and convincingly argues that the effect of sentiment can be partly attributed to the systematic heterogeneities at the household level. Note that each component of ICS/CCI measures how widely the specific subjective feeling is diffused throughout the economy, and this asymmetry is measured by the balance between the two extreme responses. Pesaran (1987) has shown that diffusion indices, such as those that form the basis of the sentiment measures, will net out all individual heterogeneities and capture the mean of the underlying distribution provided the cross-sectional distribution of expectations is homogeneous and symmetric.

DOI: 10.1002/jae

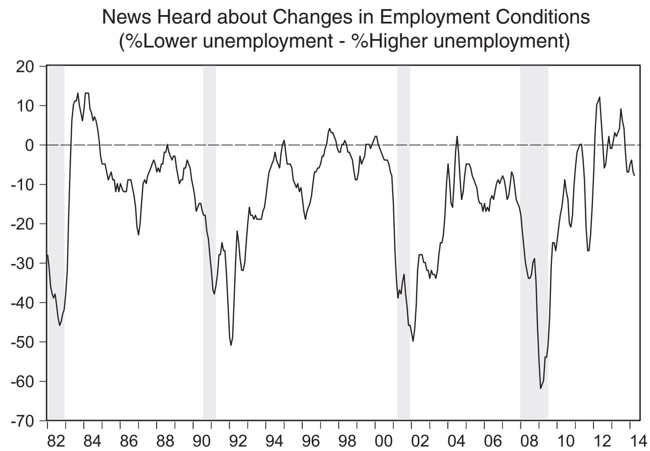


Figure 3. News heard about recent changes in employment conditions. This figure shows the news heard about recent changes in employment conditions from the University of Michigan sentiment survey. The figure plots the percentage of survey respondents who report to have heard favorable news minus the percentage having heard unfavorable news about employment conditions in the past few months prior to the survey month. Source: Chart 24b: News Heard About Change in Employment; http://www.sca.isr.umich.edu/charts.php, retrieved 11 November 2014

However, there is substantial evidence that these distributions are not uniform across households. Specifically, Souleles (2004) has shown that aggregate shocks hit different population groups differently and create substantial skewness in response patterns. In addition, to illustrate how the same economic news affects households differently, in Figure 3 we have plotted the difference between the percentage of households reporting having heard of favorable and unfavorable economic news regarding employment growth in the past few months. Bad employment news is seen to be registered and recalled overwhelmingly more than good news, and even during periods of extended high employment like the 1990s households seem not to internalize good news. The diffusion index is negative most of the times, and is typical of many of the other diffusion indices in CAB. As highlighted by Souleles (2004) and McGranahan and Toussaint-Comeau (2006), households with lower socio-economic status (less educated, low income, minority, etc.), who comprise a major part of the CAB sample, continue to hear bad economic news even during times of relative prosperity. Thus the diffusion indices that are constructed from qualitative responses reflect at least part of the skewness and heterogeneity in the cross-sectional distributions. This is one of the reasons why researchers have found that the critical threshold value (above which the economy is associated with expansion) for the well-known PMI diffusion index is significantly less than the theoretical value of 50 (see Koenig 2002; Lahiri and Monokroussos, 2013). The resulting aggregate measures of sentiment therefore incorporate the asymmetry in the cross-sectional distributions that is not fully reflected in standard macroeconomic variables. As Cochrane (1994) points out, the sentiment indices successfully aggregate idiosyncratic information from many sources and individuals.

Regarding hypothesis (ii), Lahiri and Zhao (2014) provide comprehensive evidence on the information content of consumer sentiment at the micro level. Using household data from CAB during 1978:1–2014:8, they show that sentiment captures predominantly household-specific perceptions and expectations of their own economic conditions as well as the condition and outlook of the economy. As Ludvigson (2004) has noted, higher confidence levels can be related to higher future consumption if a proportion of households are liquidity constrained. In the same CAB survey, the following question was asked in 20 selected months during 1973–1979 that identifies the liquidity-constrained

DOI: 10.1002/jae

households:[16] 'Thinking of your financial situation just now, do you feel you are in an especially good position to buy some of the things you would like to have, or is now a rather bad time for you to spend money or what?' The survey results show that an overwhelming 55.4–68.5% of the households responded by saying that they 'want to buy but can't' due to 'lack of money or extra funds' (response category 5).[17] Even though one may argue whether this group truly represents the liquidity-constrained households in a strict economic sense, it is reasonable to assume that they will respond to transitory movements in current and expected future incomes.

4. CONCLUSION

In this paper, we re-examine the role of consumer confidence surveys in forecasting personal consumption expenditure. Existing models in the literature rely mostly on simple regressions and are limited in terms of data frequency, data vintages, and number of predictors used. So, in a first step, we revisit and extend these models using both quarterly and monthly data, both in real time and using revised data vintages. Our exercises provide concrete evidence, in a more realistic and general context, of the notable contribution of confidence measures to the in-sample fit of personal consumption expenditure. However, we do not consider these relatively simple models to be sufficiently robust, even with the use of real-time data and monthly consumption expenditures. An important issue is the limited amount of information these models are based on, which make them unlikely to be fair representations of what information households use in practice when making real-time consumption decisions at multiple horizons.

We thus further consider the ability of confidence to forecast consumption in an even more realistic setting that accounts for data frequency and vintage issues in a rich information context. We use a dynamic factor model with a real-time jagged-edge dataset of over 160 explanatory variables. In this framework, we first examine the effect of the whole confidence series on the in-sample fit. Then, we examine the contribution of the latest release of consumer confidence measures on out-of-sample forecast accuracy, accounting for varying release schedules and publication lags for all explanatory variables. Next, we perform our most realistic forecasting exercise, which helps to unveil the real-time effect of consumer confidence on consumption. The results from our analysis establish that measures of consumer confidence in general make a notable and positive contribution to forecasting personal consumption expenditure. This link manifests largely through spending on services. Moreover, the effect of sentiment is found to be stronger during the last recession when all categories of consumption seem to have been affected. We have argued that the sentiment diffusion indices partially reflect the asymmetry in the way households process information from heterogeneous sources, and this may very well be one of the factors making sentiment significant in time series forecasting of consumption growth. Binding liquidity constraints faced by some households may be another factor.

ACKNOWLEDGEMENTS

The authors thank the seminar participants at Birkbeck College, Magyar Nemzeti Bank, the 9th Annual CIRANO-CIREQ Workshop on Data Revision in Macroeconomic Forecasting as well as Dean Croushore, Domenico Giannone, Massimiliano Marcellino, the editor and two anonymous referees of

[16] The question, which has largely escaped the attention of economists, was included in the following survey months: May and Nov 1973, August 1974, February, May and August 1975, February August and November 1976, February, May, August and November 1977, February May, August and November 1978, as well as February, August and November 1979.

[17] Responses were collected in the following categories: (1) good time; good time qualified; not bad, always a good time; as good a time as any; (2) Good time but don't want anything, or won't buy anything; (3) pro–con (good in some ways, bad in others); always a bad time but now as good as any; (4) bad time but can buy, or will buy; (5) bad time; bad time qualified, not good, always a bad time, want to buy but can't, lack of money or extra funds; (6) if we need (or want it) we buy it, able to buy (use only if not codeable in 1–5); (7) Have no wants or needs; old; don't buy much (use only if not codeable in 1–5); (8) don't know; (9) NA.

DOI: 10.1002/jae

the journal for many helpful comments and suggestions. Generous help from Kyle Brown, Richard Curtin, Ataman Ozyildirim, and David Small is acknowledged regarding the datasets used in this study. The content of this study does not reflect the official opinion of the European Union. Responsibility for the information and views expressed in the study lies entirely with the authors.

REFERENCES

Akerlof GA, Shiller RJ. 2010. *Animal Spirits*. Princeton University Press: Princeton, NJ; 1–7.

Banbura M, Giannone D, Modugno M, Reichlin L. 2013. Now-casting and the real-time data-flow. In *Handbook of Economic Forecasting*, Elliott G, Timmermann A (eds), Vol. 2. Elsevier: Amsterdam; 195–236.

Banerjee A, Marcellino M, Masten I. 2008. Forecasting macroeconomic variables using diffusion indexes in short samples with structural change. In *Forecasting in the Presence of Structural Breaks and Model Uncertainty*, Rapach DE, Wohar ME (eds), Vol. 3. Emerald Group: Singapore; 149–194.

Barsky R, Sims E. 2012. Information, animal spirits, and the meaning of innovations in consumer confidence. *American Economic Review* **102**: 1343–1377.

Bram J, Ludvigson S. 1998. *Does consumer confidence forecast household expenditure? A sentiment index horse race*. Federal Reserve Bank of New York: New York.

Carroll CD, Fuhrer JC, Wilcox DW. 1994. Does consumer sentiment forecast household spending? If so, why? *American Economic Review* **84**: 1397–1408.

Clark TE, McCracken MW. 2011. *Oxford Handbook on Economic Forecasting* Clements MP, Hendry DF (eds). Oxford University Press: New York; 415–440.

Cochrane JH. 1994. Shocks. *Carnegie–Rochester Conference Series on Public Policy*, Vol. 41: Atlanta; 295–364.

Croushore D. 2005. Do consumer confidence indexes help forecast consumer spending in real time *North American Journal of Economics and Finance* **16**: 435–450.

Croushore D, Stark T. 2001. A real-time data set for macroeconomists. *Journal of Econometrics* **105**: 111–130.

Curtin R. 2007. Consumer sentiment surveys: worldwide review and assessment. *Journal of Business Cycle Measurement and Analysis* **1**: 7–42.

Diebold FX, Mariano RS. 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* **13**: 253–263.

Giacomini R, Rossi B. 2010. Forecast comparisons in unstable environments. *Journal of Applied Econometrics* **25**: 595–620.

Giannone D, Reichlin L, Small D. 2008. Nowcasting: the real-time informational content of macroeconomic data. *Journal of Monetary Economics* **55**: 665–676.

Giannone D, Modugno M, Reichlin L, Small D. 2010. Nowcasting in real-time, *16th International Conference on Computing in Economics and Finance*. Society for Computational Economics: London. 13 July.

Harvey A, Leybourne SJ, Newbold P. 1997. Testing the equality of prediction mean squared errors. *International Journal of Forecasting* **13**: 281–291.

Howrey EP. 2001. The predictive power of the index of consumer sentiment. *Brookings Papers on Economic Activity* **1**: 175–216.

Ivanova D, Lahiri K. 2001. When should we care about consumer sentiment? Evidence from linear and Markov-switching models. *Indian Economic Review* **36**: 153–169.

Katona G. 1951. *Psychological Analysis of Economic Behavior*. McGraw-Hill: New York.

Koenig E. 2002. Using the purchasing managers' index to assess the economy's strength and the likely direction of monetary policy. *Federal Reserve Bank of Dallas Economic and Financial Policy Review* **6**: 1–14.

Lahiri K, Monokroussos G. 2013. Nowcasting US GDP: the role of ISM business surveys. *International Journal of Forecasting* **29**: 644–658.

Lahiri K, Zhao Y. 2014. *Determinants of consumer sentiment: evidence from household survey data Department of Economics working paper*. University at Albany: SUNY.

Ludvigson S. 2004. Consumer confidence and consumer spending. *Journal of Economic Perspectives* **18**: 29–50.

McGranahan L, Toussaint-Comeau M. 2006. Variations in consumer sentiment across demographic groups. *Federal Reserve Bank of Chicago Economic Perspectives* **30**: 19–38.

Orphanides A. 2001. Monetary policy rules based on real-time data. *American Economic Review* **91**: 964–985.

Pesaran MH. 1987. *The Limits to Rational Expectations*. Basil Blackwell: Oxford.

Rossi B. 2013. Advances in forecasting under instability. In *Handbook of Economic Forecasting*, Part B, Elliott G, Timmermann A (eds), Vol. 2. Elsevier: Amsterdam; 1203–1324.

Souleles NS. 2004. Expectations, heterogeneous forecast errors, and consumption: micro evidence from the Michigan Consumer Sentiment Surveys. *Journal of Money, Credit and Banking* **36**: 39–72.

Stock JH, Watson MW. 2011. Dynamic factor models. In *Oxford Handbook on Economic Forecasting*, Clements MP, Hendry DF (eds). Oxford University Press: Oxford; 35–60.

DOI: 10.1002/jae