Comparative Performances of Measures of Consumer and Economic Sentiment in Forecasting Consumption: A Multi-Country Analysis

Robert Gausden

Economics and Finance, University of Portsmouth, Portsmouth, U.K.

Mohammad S. Hasan

Kent Business School, University of Kent, Canterbury, U.K.

Corresponding Author:

Dr Robert Gausden, Economics and Finance, Portsmouth Business School, Faculty of Business and Law, University of Portsmouth, Portland Street, Portsmouth, PO1 3DE, U.K.

Telephone Number: +44(0)2392844031

E-mail Address: Robert.Gausden@port.ac.uk

ORCID: 0000-0003-2013-9504

This paper seeks to compare the capabilities of assorted measures of consumer and economic sentiment in predicting the growth of household expenditure. An analysis of quarterly data on five European countries shows that for none of these can the model which incorporates the EU's headline consumer confidence indicator be deemed to be significantly inferior to any of its seven rivals. However, the rankings of the sentiment variables are seen to be influenced by: the proportion of total spending by households that is devoted to durable goods; and the nature of the behaviour of consumption over the forecast interval.

Keywords: consumption; consumer confidence; model confidence set; consumer survey; forecasting.

Subject classification codes: E21; E27

I. Introduction

With regard to western industrialised countries, household final consumption expenditure represents the largest component of aggregate demand. For example, over the course of the ten years from 2007 to 2016, in the US, consumer spending averaged 68 per cent of GDP. The UK possesses the largest figure out of the major European nations, with the

corresponding mean value being equal to 67 per cent. Consequently, for this collection of countries, it would seem that general economic performance is heavily influenced by the behaviour of the personal sector. Indeed, as a form of confirmation, in a speech that was delivered at the London School of Economics on 16th January 2017, the Governor of the Bank of England, Mark Carney, referred to the growth in demand that had occurred in the UK during Autumn 2016 as being increasingly consumption-led. The evidence suggested that households were "entirely looking through Brexit-related uncertainties".[1]

The close connection that exists between the paths of consumer spending and GDP suggests that the ability of policy-makers to manage successfully the national economy would be aided by the production of accurate forecasts of household expenditure. While the Life-Cycle and Permanent Income theories maintain that consumption is governed by the total (human and non-human) resources which are available to a household, various studies have sought to investigate whether there is an additional explanatory role for psychological factors. An example is the research that was conducted by Carroll, Fuhrer and Wilcox (1994). In particular, the latter examined the forecasting performance of the University of Michigan's Index of Consumer Sentiment (ICS) in respect of the growth of total real personal consumption expenditure in the US, as well as three more specific categories of spending. A tentative conclusion that was reached by Carroll, Fuhrer and Wilcox (1994) was that data on consumer confidence incorporate information which is of relevance for future changes in consumption, independent of that which is contained in standard macroeconomic or financial indicators.[2, 3]

[1] The speech, entitled "Lambda", can be accessed online from the website of the Bank of England (www.bankofengland.co.uk/publications/Pages/speeches).

[2] In a more recent paper, Barsky and Sims (2012) were able to corroborate the finding of Carroll, Fuhrer and Wilcox (1994), that consumer confidence Granger-causes consumption, operating in the context of a three-variable vector autoregressive model.

[3] Starr (2012) has also investigated the relationship between movements in the ICS and aggregate fluctuations in the US. Founded upon a seven-variable structural vector autoregressive model, impulse responses were generated which enabled the conclusion to be reached that, even after having extracted the effects of news

In contrast to Carroll, Fuhrer and Wilcox (1994), Acemoglu and Scott (1994) analysed data on the UK. More specifically, in respect of the quarterly growth of *per capita* household expenditure on non-durable goods, they explored the predictive content of a measure of consumer sentiment that was founded upon a survey which was undertaken by Gallup and commissioned by the European Union.[4] Estimation results showed that the confidence variable (lagged one or two quarters) was significantly related to the current change in consumption. Such a finding was in contradiction of the Rational Expectations-Permanent Income Hypothesis, which Acemoglu and Scott (1994) explained in terms of a desire for precautionary saving by households, rather than capital market imperfections.

Subsequently, various studies have sought to appraise the forecasting performances of different elements of the respective aggregate sentiment indicator, as well as the overall measure. For example, in a wide-ranging paper, Bram and Ludvigson (1998) conducted an assessment of the predictive accuracy of the University of Michigan's ICS and the Conference Board's Consumer Confidence Index, as well as their five components. In the context of forecasting the growth of five categories of personal spending in the US, the Conference Board's gauge of sentiment appeared to be generally more proficient than the ICS. Further analysis showed that reflection on how the financial position of the respondent's family had changed over the past twelve months, as was requested in the University of Michigan's survey, was unhelpful, while answers to questions about the availability of jobs, which featured in the Conference Board's survey, tended to improve the quality of predictions.

Operating in conjunction with quarterly data on the UK, Easaw and Heravi (2004) adopted a quantitative approach that was similar to that of Bram and Ludvigson (1998). The

---

shocks, innovations in consumer sentiment still have a significant, positive influence on personal consumption expenditure.

[4] The design of the sentiment variable accords with the construction of the GfK indicator, which will be mentioned later in this paper.

measure of consumer confidence that was central to this empirical exercise was the GfK index. However, consideration was given to the forecasting performance of not only the aggregate variable but also each of its two expectations components and a summary indicator of whether there are believed to be benefits to households from undertaking major purchases at the present time. Regression functions were constructed to explain the growth of five different categories of consumption expenditure. An evaluation of the accuracy of one-step-ahead recursive predictions showed that data on consumer sentiment were generally of greater use when attitudes were more sanguine.[5] However, the views that were held by individuals concerning the timing of large-scale acquisitions could be regarded as being of greater relevance during unfavourable economic conditions.

In a multi-country study,[6] Cotsomitis and Kwan (2006) elected to assess the forecasting capabilities of the EU's headline Consumer Confidence Indicator (CCI), together with its four constituent parts. Additionally, attention was directed towards the predictive performance of the EU's more broad-ranging Economic Sentiment Indicator (ESI).[7] On the basis of a within-sample analysis that was designed to account for the growth of real total consumption expenditure, although results were somewhat diverse, there was a suggestion that, in general, the incremental explanatory power of the ESI exceeded that of the CCI.

Also, employing seven different confidence measures, which included an EU aggregate CCI, one-quarter-ahead recursive forecasts of the growth of total consumption were produced over the interval, 1999q1 – 2002q3. In only a distinct minority of instances were the sentiment data found to enhance the accuracy of the predictions, encouraging

[5] Especially data on the aggregate GfK index.
[6] More specifically, empirical analysis was conducted in conjunction with quarterly data on nine EU countries.
[7] As will be clarified in the next section, the ESI represents a composite measure of sentiment, which is formed from a combination of five sectoral indicators.

Cotsomitis and Kwan (2006) to draw the conclusion that the respective confidence indices offered limited information about the future behaviour of household spending.[8]

Gausden and Hasan (2016) also sought to establish the benefits from using data on the EU's CCI in forecasting, but restricted their study to solely the UK. Out-of-sample predictions were generated of the quarterly growth of five different categories of household final consumption expenditure using regression models which incorporated lags on the headline indicator, as well as two modified versions of the latter. The adaptations to the CCI had the general consequence of increasing forecast accuracy over the interval, 2008q1 – 2013q1. However, overall, any gains that were achieved from using sentiment data were found to be confined to the recent period of economic crisis.

On the basis of the literature that has just been reviewed, it is apparent that different characterisations of consumer and economic confidence have been employed for the purpose of generating predictions of the growth of personal spending. Consequently, the principal objective of this paper is to compare the forecasting capabilities of a number of these indicators, with special attention being paid to the relative performance of the EU's headline measure, the CCI. Regarding the earlier investigations, where a formal out-of-sample evaluation has been undertaken, this has consisted of the application of tests that have been proposed by Clark and McCracken (2002) and Harvey, Leybourne and Newbold (HLN) (1997). The respective procedures seek to examine whether or not a *pair* of series of predictions exhibit the same degree of accuracy. A distinguishing aspect of the current study is that *additionally* it adopts a statistical approach which suits a situation in which there are several candidate forecasting models. In particular, Hansen, Lunde and Nason (2011) have recommended a sequential method involving a chosen loss function and an elimination rule, which is designed to reduce an initial group of objects to a smaller admissible collection. The

[8] Possibly, the largely negative findings were influenced by the short length of the post-sample data period.

latter is referred to as the Model Confidence Set (MCS) and can be regarded as including the best specification with a stated probability. A feature of the technique that has been advocated by Hansen, Lunde and Nason (2011) is that if the data are insufficiently informative then few, if any, of the original equations will be excluded from the MCS. For this reason, it is decided to conduct an econometric exercise which uses data on as many as five principal European countries (France, Germany, Italy, the Netherlands and the UK).[9] As a result of employing the MCS procedure, based upon probability values, a ranking can be achieved of the models which comprise the full set. Through relating the standings of the relevant functions to the characteristics of the countries' consumption data, it is possible to establish conditions under which a reliance upon the CCI for forecasting is comparatively productive.

Within this study, for each of the five European countries, eight different regression equations are assembled, featuring seven different representations of consumer confidence. Estimation is undertaken using quarterly data in conjunction with a rolling sample period, which gives rise to the generation of thirty one-step-ahead predictions of the growth of consumption expenditure over the interval, 2008q4 – 2016q1. Following application of MSE-t tests at the five per cent level, the model incorporating the CCI is never found to be significantly outperformed by any of its competitors. Also, for none of the five European countries is the EU's CCI excluded from the ninety per cent MCS. Hence, the statistical evidence can be interpreted as justification for the prominence that has been given to this indicator.

However, in the context of the MCS analysis, across the countries, differences can be observed in terms of the number of surviving models and the rankings of the regression

[9] The analysis is restricted to involving European countries for the reason that consistency can be achieved in terms of the sentiment data which are obtained from regular harmonised surveys that have been undertaken by the Directorate General for Economic and Financial Affairs of the European Commission.

functions. For Italy and the Netherlands, the greater variation that is exhibited by the consumption data over the forecast interval offers more scope for discriminating between competing processes. However, the orderings of the equations are not the same for these two countries. More specifically, in the case of Italy, the relatively small proportion of total consumer spending that is devoted to durable goods renders the CCI more relevant than the GfK index. With regard to the Netherlands, though, an indicator of sentiment that is based purely upon answers to forward-looking questions is seen to be unhelpful during a period of generally sluggish consumption behaviour.

This paper has four remaining parts. In Section II, details are provided of the confidence measures that feature in the empirical analysis. Section III is largely dedicated to an outline of the MCS methodology that is fundamental to this study, but also discusses the use of the MSE-t test. Section IV contains a description of the time series on consumption and a summary of the sentiment data, which is followed by the presentation and a discussion of the econometric results. Finally, concluding remarks and recommendations are made in Section V.

II. Consumer and Business Confidence Indicators

This section introduces the seven consumer and business confidence indicators which enter the subsequent empirical analysis. For every one of the sentiment measures, the respective (quarterly, seasonally-adjusted) data series is obtained from the Directorate General for Economic and Financial Affairs (DGEFA) of the European Commission. The DGEFA regularly conducts monthly surveys of six different sectors of countries that currently belong to or are seeking to join the EU. In particular, information is gathered in relation to the manufacturing and construction industries, consumers, the retail trade, services and, more specifically, financial services.

The objective of the monthly consumer survey is to acquire, through both direct and indirect means, an indication of households' spending and savings intentions. A sample of individuals are presented with twelve qualitative questions concerning the financial position of their household, the national economic situation, developments in consumer prices, changes in the number of unemployed, the timing of and the amount that is spent on major purchases, and the attractiveness and likelihood of saving.[10] For each of the questions, an aggregate balance (B) is created essentially by calculating the difference between the numbers of positive and negative responses and expressing the result as a percentage of the total number of answers.[11]

The EU's headline measure of household sentiment is the CCI. The latter is formed from the balances relating to merely four of the twelve questions which are the foundation of the consumer survey.[12] In each case, the respondent is asked to give consideration to either individual financial or macroeconomic circumstances twelve months' ahead. The specific questions are detailed below.

Question 2

How do you expect the financial position of your household to change over the next 12 months?

Question 4

How do you expect the general economic situation in this country to develop over the next 12 months?

Question 7

How do you expect the number of people unemployed in this country to change over the next 12 months?

[10] All twelve questions are presented in Table A2 in the appendix to this paper.
[11] Hence, B has the potential to range from -100 to +100.
[12] The questions are selected with the aim of maximising the coincident correlation between the CCI and the reference variable (i.e., the growth of household consumption).

Question 11

Over the next 12 months, how likely is it that you save any money?

If the aggregate balances corresponding to these four questions are denoted by Q2, Q4, Q7 and Q11 then the value of the CCI is achieved through calculating a simple arithmetic average, i.e.,

$$CCI = \frac{1}{4}(Q2 + Q4 + Q7 + Q11)$$

In a recent paper, Gausden and Hasan (2016) proposed a minor modification to the CCI. Reasons were supplied for seeking to deny Q11 a contribution towards the creation of an aggregate measure of consumer sentiment.[13] Hence, there was advocated the alternative construct:

$$CCI^{*} = \frac{1}{3}(Q2 + Q4 + Q7)$$

Subsequently, empirical support was received for the discarding of Q11. In particular, for each of five different categories of UK household consumption expenditure, with reference to the forecast interval, 2008q1 - 2013q1, more accurate one-quarter-ahead predictions were generated by virtue of including CCI^{*}, rather than CCI, in the respective regression model.[14]

The barometer of household confidence that is favoured by the GfK organisation combines the balances corresponding to five of the questions that provide the basis of the European Commission's consumer survey. In contrast to the CCI, the GfK indicator does not depend upon solely responses to forward-looking questions. More specifically, in addition to

[13] Specifically, see p. 1704 of the article by Gausden and Hasan (2016).
[14] See Table 8, p.1706, within the paper by Gausden and Hasan (2016).

relying upon the net scores that are derived from questions 2 and 4, use is made of the answers to the following three questions:

Question 1

How has the financial situation of your household changed over the last 12 months?

Question 3

How do you think the general economic situation in the country has changed over the past 12 months?

Question 8

In view of the general economic situation, do you think that now is the right moment for people to make major purchases such as furniture, electrical/electronic devices etc.?

If the balances that are associated with the above three questions are denoted by Q1, Q3 and Q8, respectively, then the GfK aggregate measure of consumer sentiment can be defined as:

$$GfK = \frac{1}{5}(Q1 + Q2 + Q3 + Q4 + Q8)$$

The GfK indicator has featured in empirical research that has been conducted by, *inter alia*, Easaw and Heravi (2004), Nahuis and Jansen (2004), Easaw, Garratt and Heravi (2005) and Gausden and Hasan (2012). In particular, in the study by Easaw and Heravi (2004), in respect of five different categories of UK household consumption expenditure, an evaluation was performed of the post-sample predictive capabilities of not only GfK, itself, but also three of its components, namely, Q2, Q4 and Q8. The current paper will follow this practice, although is restricted to forecasting solely the growth of total spending by the personal sector.

As was mentioned in the introduction, the multi-country analysis that was undertaken by Cotsomitis and Kwan (2006) sought to examine the explanatory and predictive power of not only the CCI and its constituent elements but also the more general confidence measure, the ESI. The latter was designed in 1985 by the European Commission with the intention of reflecting developments in not only the consumer sector but also industry, construction, the retail trade and services.[15] For each of these categories of the economy, a survey is conducted by the DGEFA, which permits an overall balance to be produced that summarises the attitudes and assessments of the sample members. For the purpose of calculating the value of the ESI, the individual aggregates are allocated weights of 0.4 (industry), 0.3 (services), 0.2 (consumers), and 0.05 (each of construction and retail trade).[16, 17] It should be noted that, unlike all of the other confidence indicators to which reference has been made, the ESI is deliberately fashioned such that it has a long-term mean of 100 and standard deviation of 10.

The within-sample results that were reported in the paper by Cotsomitis and Kwan (2006) showed the ESI to be generally a more useful explanatory variable than the CCI.[18] Additionally, irrespective of the predominantly poor post-sample forecast performances of all of the sentiment measures that entered this study, the former could, once again, be regarded as superior to the latter.[19] Hence, in spite of its relatively broad emphasis, there appears to be justification for permitting the inclusion of the ESI in the group of rival predictors of the growth of consumer spending.

[15] Hence, the focus of this measure is future changes in GDP, rather than, more specifically, household expenditure.
[16] The allocations are decided by two factors: the representativeness of the sector; the ability of the variable to track the growth of GDP. To be more precise, the weights are applied to the standardised forms of the components of the respective aggregates.
[17] In contrast, within the study by Cotsomitis and Kwan (2006), the proportions consisted of 0.4 (industry) and 0.2 (each of consumers, construction and retail trade).
[18] See Table 3 (p. 603) and Table 4 (p. 604).
[19] See Table 6 (p. 607).

III. Empirical Methodology

Within this study, two types of regression model are constructed for the purpose of forecasting the quarterly change in the logarithm of household consumption expenditure:

$$\Delta \ln(Cons_t) = A_0 + \sum_{j=1}^{n_1} A_j \Delta \ln(Cons_{t-j}) + u_t ; \quad (1)$$

$$\Delta \ln(Cons_t) = A_0 + \sum_{j=1}^{n_2} A_j \Delta \ln(Cons_{t-j}) + \sum_{j=1}^{n_3} B_j S_{t-j} + u_t. \quad (2)$$

With regard to these two specifications, $Cons$ constitutes an abbreviation of total household final consumption expenditure, while S denotes the measure of consumer or economic sentiment, which is represented, in turn, by CCI, CCI*, GfK, Q2, Q4, Q8 and ESI. Also, u_t is signifying the customary stochastic disturbance term. Equation (1) can be interpreted as a form of baseline equation, which simply allows the current value of $\Delta \ln(Cons)$ to be related systematically to its own past. In contrast, equation (2) permits the growth of consumption in the present period to be dependent upon lagged values of both itself and the relevant confidence indicator. The maximum length of lag, n_1, n_2 or n_3, is always constrained to be 4 quarters, while the minimum value of n_3 is set at 2.[20] Operating within these bounds, optimal values of n_1, n_2 and n_3 are decided by the Schwarz information criterion, having estimated all equations by Ordinary Least Squares over a sample period that extends to 2008q3.[21, 22]

[20] This accords with the practice that has been adopted by the authors when undertaking earlier research in this area.

[21] As will be explained in the next section of the paper, the start date of the estimation period is not the same for all five of the countries.

[22] It should be clarified that, with respect to both equation (1) and equation (2), the possibility is entertained of no lags being merited on the consumption variable. Indeed, in the absence of $\Delta \ln(Cons_{t-j})$ on the right-hand side of equation (1), the regression function conforms to the Rational Expectations-Permanent Income Hypothesis.

For all five of the countries, one-step-ahead forecasts of $\Delta\ln(Cons_t)$ are generated over the interval, $t = 2008q4, 2009q1, \dots, 2016q1$, founded upon a rolling regression procedure. More specifically, the prediction for 2008q4 rests upon an equation that has been estimated using data through to 2008q3. The subsequent (twenty-nine) forecasts are achieved by continually moving the start and end dates of the sample period forward by one quarter. Hence, for each country, there are produced eight sets of forecasts of $\Delta\ln(Cons_t), t = 2008q4, 2009q1, \dots, 2016q1$, corresponding to equation (1) and the seven variants of equation (2). By combining the predictions with the respective actual values of the dependent variable, eight series of forecast errors can be obtained. The predictive accuracy of the eight models can subsequently be compared by calculating the values of a summary statistic, such as the mean square error (MSE). It should be appreciated, though, that concern is not simply with whether there are any differences in the values of the MSE. More fundamentally, the focus is upon whether any disparities which are identified are statistically significant.[23]

Within this paper, a formal econometric analysis is conducted through both implementing the MCS procedure and applying (pairwise) tests which seek to examine whether or not two rival series of forecasts are of identical quality. For the purpose of introducing these statistical methods, the assumption is made that there are n realised values of a variable, y, that are denoted by $y_1, y_2, \dots, y_n$. The corresponding one-step-ahead predictions that are produced by model i are represented by $\hat{y}_{i,1}, \hat{y}_{i,2}, \dots, \hat{y}_{i,n}$. Through a comparison of the respective actual and forecast values, it is possible to create a loss function, $L_{i,t}, t = 1, 2, \dots, n$, which, in this study, takes the form of the square of the prediction error. Should a series of forecasts be generated by each of two rival functions, i and j, then there is the potential to construct a loss differential, $d_{ij,t} = L_{i,t} - L_{j,t}, t = 1, 2, \dots, n$. On the basis of its design, a positive (negative) value of the latter is an indication that model i is responsible

[23] Diebold and Mariano (1995) have observed the tendency in the literature for relevant point estimates to be produced without consideration being given to their sampling uncertainty.

for a less (more) accurate prediction than model j. The arithmetic mean, $\bar{d}_{ij} \equiv \frac{1}{n}\sum_{t=1}^{n} d_{ij,t}$ provides a summary measure of the relative merits of the two competing equations, i and j.

Consultation of the review chapter by Clark and McCracken (2013) encourages reliance upon the MSE-t test for the purpose of appraising the null hypothesis of equal predictive ability of two rival series of forecasts. The definition of the MSE-t statistic, which is shown below, corresponds to that which can be found in Table 20.1 of the aforementioned publication:

$$MSE - t = \frac{n^{-1/2} \sum_{t=1}^{n} d_{ij,t}}{\hat{S}_{dd}^{1/2}}$$

where $\hat{S}_{dd}^{1/2}$ represents the estimate of the long-run variance of $d_{ij,t}$.[24]

Clark and McCracken (2013) indicate the sensitivity of the distribution of the MSE-t statistic to the context in which the test is being performed. For example, they distinguish between: population-level and finite-sample evaluation; non-nested and nested models; recursive and rolling or fixed-period estimation windows. With respect to the subsequent empirical analysis, given a desire to undertake finite-sample inference in relation to nested, as well as non-nested models, and to generate one-step-ahead predictions that are founded upon a rolling estimation sample, the theory of Giacomini and White (2006) seems to be appropriate.

In contrast to Diebold and Mariano (1995) and West (1996), Giacomini and White (2006) sought to allow for estimation uncertainty when comparing forecasts. Hence, loss functions were made to depend upon estimates, rather than the true values of population parameters. Also, a feature of their paper was an acceptance that the forecasting model may be subject to misspecification. Under such a circumstance, it was often found to be beneficial

[24] The expression for the MSE-t statistic has been adapted to accord with the notation that is used in this paper.

to use a predictor that was based upon a limited, rather than an expanding memory, leading to a preference for generating rolling window forecasts. Finally, Giacomini and White (2006) provided a framework that permitted a unified treatment of nested and non-nested models and possessed the advantage of allowing the computed value of the test statistic to be obtained using standard econometric software. Indeed, in section 3.4 (p. 1557) of their paper, there is presented what amounts to an MSE-t or Diebold-Mariano test statistic that can be employed in conjunction with a standard normal distribution.[25]

In this paper, for the purpose of performing MSE-t tests, the theory of Giacomini and White (2006) is combined with the practice of Clark and McCracken (2013).[26] Hence, the estimate of the long-run variance is produced, having chosen a rectangular kernel and a bandwidth, which is equal to the forecast horizon less one. Additionally, $\hat{S}_{dd}^{1/2}$ is subject to the finite-sample adjustment that has been proposed by Harvey, Leybourne and Newbold (1997). Subsequently, computed values of the MSE-t statistic are contrasted with critical values corresponding to a standard normal distribution.

The application of an MSE-t test can be regarded as somewhat limited as this merely allows pairwise comparisons of forecasting performance. Granted that equations (1) and (2) incorporate eight distinct regression functions that will be estimated for the purpose of predicting the growth of consumption, it seems to be appropriate to utilise additionally a more broad-ranging statistical technique. In particular, the MCS procedure, which is attributable to Hansen, Lunde and Nason (2011), is suitable for a situation in which there are several candidate equations. Through the implementation of sequential testing in conjunction with an elimination rule, the aim is to reduce an original group of objects to a smaller set – known as the MCS – which contains the best specification with a stated probability. Reliance

[25] The asymptotic theory of Giacomini and White (2006) requires the size of the estimation sample to be small, relative to the number of post-sample observations. However, Clark and McCracken (2013) maintain Monte Carlo evidence is limited on the respective magnitudes that are necessary for accurate inference.
[26] In particular, see section 3.3 in the chapter by Clark and McCracken (2013).

upon this empirical approach appears to be especially relevant in the current study if consideration is being given to whether the CCI merits a continuation of its status as the EU's headline measure of consumer confidence.

More formally and in greater detail, application of the MCS procedure assumes that there are m rival functions that are contained within the set, M_0, which are used to generate predictions of the future values of a common variable, e.g., $y_1, y_2, \dots, y_n$. Once again, the average loss differential, $\bar{d}_{ij}$, can be calculated to provide a summary of the relative forecasting performances of two competing models, i and j. In this context, though, in relation to an individual function, i, within M_0, there is the potential to make $m-1$ comparisons, and so to produce the same number of values of the mean differential, i.e., $\bar{d}_{ij}$, $j = 1, 2, \dots\dots\dots, m$ $(j \neq i)$. The average of $\bar{d}_{ij}$ across j yields $\bar{d}_{i.}$, the standardised version of which is shown below:

$$t_{i.} = \frac{\bar{d}_{i.}}{\sqrt{\widehat{var.(\bar{d}_{i.})}}}$$

where $\widehat{var.(\bar{d}_{i.})}$ denotes a suitable estimator of the population variance of $\bar{d}_i$.

This form of statistic can be produced for every one of the m models within M_0. It is then possible to assemble from $t_{i.}$ $(i = 1, 2, \dots\dots\dots, m)$,

$$T_{max,M} = max.\, t_{i.}, \text{ where } i \in M_0,$$

that has a non-standard asymptotic distribution, which is consequently estimated using a bootstrap procedure. A significant value of $T_{max,M}$ necessitates a model to be discarded from

the set. The obvious elimination rule requires a removal of the object which is associated with the maximum value of $t_{i.}$ $(i = 1, 2, \ldots\ldots\ldots, m)$. The outlined procedure should then be repeated, but with the set containing one less model than before. Sequential testing should continue until the respective null hypothesis cannot be rejected.[27] The surviving elements form the MCS.

Within the relevant literature, the 100(1 – α) per cent MCS tends to be denoted by $\widehat{M^*_{1-\alpha}}$, where α indicates a common level of statistical significance at which individual tests of equal predictive ability are performed.[28] The dimension of $\widehat{M^*_{1-\alpha}}$ is symbolised by m^*, which is necessarily ≥ 1 and $\leq m$. Corresponding to each one of the original models within M_0, a probability value ($\widehat{p_i}$, $i = 1, 2, \ldots\ldots\ldots, m$) can be calculated. Object i is permitted to enter the MCS on condition that $\widehat{p_i} \geq \alpha$. The lower is its probability value, the less desirable is a specification, and the optimal equation can always be identified through being associated with a probability of 1.

IV. Description of the Data and Empirical Results

A. Data

For the purpose of conducting the econometric analysis, quarterly, seasonally-adjusted time-series data have been obtained on France, Germany, Italy, the Netherlands and the UK. The study is restricted to involving European countries such that there is consistency in the construction of the aggregate sentiment indicators. Additionally, a harmonised approach has been adopted by the European Commission DGEFA in collecting and processing the underlying data. For each of the seven confidence measures that have been presented in section II, monthly data are available from the beginning of 1985. However, three-month

[27] The prevailing set is signified by M. The null hypothesis maintains an equal predictive ability of the models which are incorporated within M.

[28] A positive aspect of the manner of the sequential approach which is adopted is that there is no distortion to the size of the test.

averages have been calculated so that the sentiment and expenditure data are of a matching frequency.

Quarterly time series on consumer spending, in the form of constant prices, have been accessed directly from Datastream.[29] However, the period over which the data extend is not the same for all of the five countries. In particular, in spite of a common end date of 2016q1, the start date varies, consisting of 1985q1 for France, Italy and the UK, 1988q1 for the Netherlands, and 1991q1 for Germany. Furthermore, it should be noted that, in order to achieve continuous series for both Italy and the Netherlands, it was necessary to perform a splicing operation.

Figure 1 shows, in the form of separate line graphs, the time-series data on total consumption for the five European countries. The figure reveals that, in all cases, household expenditure has increased over the period as a whole. However, a distinction can be drawn between, on the one hand, France, Germany and the UK and, on the other, Italy and the Netherlands. More specifically, for each of the first three countries, by the end of the data period, spending by the personal sector had reached its highest level. In contrast, for Italy and the Netherlands, the peak was registered much earlier, in 2007q2 and 2008q2, respectively.

An examination of Figure 1 also indicates that, in France and the UK, consumption has behaved in a similar manner. To be more precise, in both countries, there occurred a prolonged period over which household expenditure rarely fell: from 1997q2 to 2007q4 in France; and from 1992q2 to 2007q4 in the UK. However, this predominantly continuous rise was followed by a sustained decrease in spending, extending from 2008q1 to 2008q4 in France and from 2008q1 to 2009q2 in the UK, after which a form of recovery took place. In contrast, in Italy, the path of consumption has been far more turbulent, with downswings being experienced at relatively frequent intervals. Indeed, consumer spending was subject to

[29] Precise details are provided in Table A1 within Appendix A.

an uninterrupted decline not only from 2007q3 to 2009q1 but also, soon after, from 2011q1 to 2013q2.[30]

With regard to Germany, the corresponding graph shows a marked slowdown in the rate of growth of consumption from 1991q1 – 2001q4 to 2002q1 - 2009q3. However, cyclical fluctuations are largely absent from the time series and the personal sector in Germany can be viewed as exhibiting resilience in response to the financial crisis. For example, although household expenditure fell sharply in 2007q4, from this date onwards, there have never been witnessed decreases in successive periods. Finally, the pattern of behaviour of personal spending in the Netherlands seems to have been in stark contrast to that in the other four countries. The associated line graph indicates a strong and continuous growth of consumption from 1993q2 to 2001q4. A local maximum was subsequently reached in 2002q4. However, over the next fifty-three quarters, the advancement of household expenditure was distinctly uneven, with an overall increase being recorded of merely 1.8 per cent.

Figure 1

Table 1

In connection with the data on the seven different measures of economic and consumer confidence, for each of the five European countries, values of summary statistics are shown in Table 1. On the basis of the information that is contained in the table, it is possible to make the following observations. First, on the whole, households have exhibited the greatest degree of optimism in the Netherlands and the least amount in France. More

[30] It would appear, then, that collectively households in Italy have enjoyed relatively limited success in achieving consumption smoothing, which is possibly explained by cross-country differences in the timing and impact of financial deregulation. See, for example, Table A2 within the article by Al-Eyd, Barrell and Davis (2009), as well as the OECD working paper by Boone, Girouard and Wanner (2001).

specifically, for four of the consumer sentiment indicators, the Netherlands is associated with the highest mean value, while France is attached to the lowest. In relation to Q8, the ordering of the countries appears to be quite different from for the other measures. The UK is the only one of the five for which the average value is positive, while the figure for Italy is substantially below that for any of the other four nations. Indeed, for Italy, over the interval, 1985q1 – 2016q1, the value of Q8 is always negative, implying a persistent unwillingness to undertake major purchases.[31]

As was indicated in section II, the ESI is designed in such a way that it has a much broader coverage than any of the consumer confidence variables. Hence, it is entirely understandable for the rankings of the five countries according to the mean values of this indicator to be considerably out of line with those that are based upon the average values of a measure of consumer sentiment. Indeed, an inspection of the relevant figures in the final row of Table 1 reveals that the Netherlands finds itself in bottom position, while, in general, the level of optimism has been highest in the UK.

Considering the contents of Table 1, from the perspective of the representations of consumer confidence, it appears that individuals are more sanguine concerning the financial position of their own household than the general economic situation twelve months in the future. For every one of the five countries, the mean value of Q2 is markedly above that of Q4. Indeed, on the basis of a comparison of the respective average values of the indicators, for all countries, Q2 occupies one of the first two places in the rankings. Interestingly, Q8 finds itself in first position for the UK, which is in complete contrast to the situation for Italy. This measure also tends to exhibit the greatest volatility over time, granted that, in four out of five cases, it possesses the largest standard deviation out of all of the sentiment variables.

[31] Please respect that these rankings are subject to variation over time. For example, over the interval, 2008q4 – 2016q1, for every one of the indicators, the mean value for Germany is greater than that for the Netherlands.

B. Empirical Results

At the beginning of this sub-section, it should be confirmed that, for all models and time periods, forecast accuracy is measured by the square of the respective prediction error. Results that were obtained from the application of the MSE-t test are shown in the tables 2A – 2E, while, for all countries, the outcome of the MCS analysis is presented in Table 3.

A study of the contents of the six tables reveals a lack of uniformity across the five countries. For France, Germany and the UK, statistically, there appears to be far greater similarity in the forecast performances of the eight regression functions than for the Netherlands and, especially, Italy. In connection with the MSE-t test, for each country, there was the scope to undertake twenty-eight pairwise comparisons of predictive attainment.[32] For Germany, the UK and France, the numbers of occasions on which the value of $MSE - t$ was significant at the five per cent level are 0, 2 and 2, respectively. Moreover, none of the equations are excluded from the 75 per cent MCS for Germany, while only one object is removed for each of the UK and France. In contrast, in tables 2C and 2D, it is possible to observe 15 and 3 significant values of the corresponding test statistic for Italy and the Netherlands. Additionally, consideration of Table 3 indicates that there is only one surviving regression equation in the 75 per cent MCS for Italy, while half of the models have been eliminated in the case of the Netherlands.

Table 2A

Table 2B

Table 2C

[32] If the number of distinct models is equal to 8 then there are 8(8 – 1)/2 possible pairs.

Table 2D

Table 2E

Table 3

A fundamental finding which emerges from viewing the results which are presented in tables 2A – 2E is that there is no instance of, at the five per cent level, a significant value of $MSE - t$ and the MSE for the CCI model exceeding that for any other process.[33] Also, the probability values which are contained within Table 3 enable the function incorporating the CCI to feature in all of the five countries' 90 per cent MCS.[34, 35] On this basis, there would seem to be a statistical justification for the CCI continuing to fulfil the role of the headline measure of consumer confidence. However, it is accepted that the case for the CCI is weaker if the MSE-t test is performed at the ten per cent level of significance and the focus is upon the 75 per cent MCS.

Closer scrutiny is now given to the statistical results for each of the five countries. It is convenient to begin with a discussion of the findings for Germany for the reason that, as has been reported above, in no situation is there a significant value of the $MSE - t$ statistic, while none of the equations are omitted from even the 75 per cent MCS. An explanation for the lack of a distinction between the predictive performances of the eight functions is the limited information content within the data on consumption. Over the forecast interval, 2008q4 – 2016q1, household expenditure in Germany rose by 7.96 per cent. Moreover, the growth in spending appeared to be rather even. In particular, while there can be identified seven (out of thirty) instances in which consumption fell, there were never experienced

[33] Such a pronouncement cannot be made about any of the other seven models.
[34] Indeed, for both Italy and the UK, the CCI can be regarded as helping to deliver the best set of predictions, as signified by the associated probability value being equal to 1.
[35] The regression functions which incorporate the CCI, Q2, Q4 and ESI enter all countries' 90 per cent MCS.

consecutive declines. The restricted variability in the data on $\Delta \ln(Cons)$ can be seen, in Table 3, to produce a very narrow range of values of the RMSPE across the eight models.

The evidence suggests that the nature of the data on the UK has also prohibited the econometric methods from being largely discerning. As mentioned above, there are only two statistically significant values of $MSE - t$. Additionally, merely one regression function (relating to Q4) is denied entry to the 75 per cent MCS, while all eight of the objects are retained within the 90 per cent group. In contrast to Germany, over the forecast interval, the UK endured a sustained fall in consumption, lasting from 2010q3 to 2011q2 (comprising four quarters and equalling 1.5 per cent). However, the more dominant feature was the virtually uninterrupted increase which extended from 2011q3 to 2016q1 (comprising nineteen quarters and equalling 10.5 per cent).[36]

From an inspection of Figure 1, over the forecast interval, household expenditure in France seemed to exhibit a greater degree of cyclical behaviour than in either Germany or the UK. More specifically, from 2008q4 to 2011q1, consumer spending in France continually rose. Thereafter, to 2014q1, a general decline occurred which amounted to 0.5 per cent. However, from this low point in 2014q1 until the end of the prediction period, in 2016q1, another uninterrupted movement upwards was witnessed. While this greater variability in consumption ought to have facilitated the task of drawing a distinction between the forecasting capabilities of the different models, Table 2A contains only two significant values of $MSE - t$. Also, an examination of Table 3 shows that only the autoregressive process is excluded from the 75 or 90 per cent MCS. A contributing factor towards the results for France is the close correspondence between the fluctuations in the indicators of confidence. Granted that there are seven different measures of sentiment included in this study, it is possible to generate twenty-one pairwise sample correlation coefficients. For the complete

[36] The single exception occurred in 2014q4.

data period, 1985q1 – 2016q1, fifteen values are in excess of 0.75. Predictably, some of the weaker linear associations involve the ESI. Furthermore, all of the correlation coefficients relating GfK to the CCI, CCI*, Q2, Q4 and Q8 are greater than 0.8.[37]

Over the forecast interval, the cyclical variation in household expenditure appears to be more pronounced in Italy than any of the other four European countries. At the beginning of the period, Italy was enduring a fall in spending which reached a low point in 2009q1. Thereafter, consumption climbed to a peak in 2010q4, after which there occurred a prolonged decline to 2013q2.[38] Subsequently, a form of recovery was staged until the end date, 2016q1, yet expenditure still remained 3.68 per cent below its value in 2008q4.

Inter alia, both Gausden and Hasan (2016) and Lahiri, Monokroussos and Zhao (2016) have identified that information on consumer confidence serves a more productive role in forecasting household spending during an economic downturn. On this basis, perhaps it should not be surprising to find that the data on Italy, compared to those on Germany and the UK, allow a clearer distinction to be drawn between the predictive capabilities of the rival models. A study of the results that are contained in Table 2C and Table 3 shows that the equation featuring the CCI produces significantly more accurate forecasts than the function which includes the GfK index, which is indeed excluded from the 90 per cent MCS. [39]

The relevant figure in the penultimate row of Table 1 indicates a relative unwillingness of households in Italy to undertake large-scale purchases. Correspondingly, the percentage of total personal spending in Italy that is devoted to durable goods is considerably lower than for the other four countries. For example, with regard to the period, 2008q4 – 2016q1, as a percentage of overall consumption by the household sector, expenditure on durable items averaged only 7.8 in Italy, in contrast to 11.1 in Germany. For the reason that spending on durable goods represents only a small share of total consumption expenditure in

[37] Detailed results are available, on request, from the corresponding author.
[38] Between 2010q4 and 2013q2, household expenditure decreased by as much as 7.16 per cent.
[39] To more than two decimal places, the respective probability value is less than 0.10.

Italy then the responses to Question 8 in the European Commission's consumer survey may be regarded as unhelpful in terms of signalling future changes in *aggregate* consumption. Indeed, it can be observed in Table 3 that the model incorporating Q8 generates by some distance the least accurate forecasts. Additionally, as a consequence of giving a weighting of 0.2 to Q8, the GfK index is found to be markedly inferior to the CCI.

The composition of the GfK measure of confidence is similar to that of the ICS, which was devised by the University of Michigan's Survey Research Center. More specifically, both are founded upon the responses to five questions, which require reflection on the past, contemplation of the future, and a consideration of whether or not it is currently an appropriate time for a person to purchase a major household item.[40] In the paper by Mishkin *et al*. (1978), the principal author agrees with the economists at the Survey Research Center that the ICS should serve as a useful explanatory variable with regard to consumption expenditure. Mishkin also sees this indicator of confidence as being particularly applicable to spending on durable goods, although not on account of the perceived discretionary nature of the latter. Instead, the close association is understood to be derived from the sentiment measure being interpreted as representing the likelihood of financial stress or uncertainty, and durable products being identified as a relatively illiquid type of personal asset.[41]

Empirical studies of the relationship between the ICS and household expenditure on durable goods have been conducted by, *inter alia*, Kumar, Leone and Gaskins (1995) and Huth, Eppright and Taube (1994). Kumar, Leone and Gaskins (1995) found that, according to conventional statistical criteria, the framework of a Bayesian Vector Autoregressive (VAR) model, incorporating the ICS, was the most effective for predicting US durable goods'

[40] Within Box A on p. 61 of the paper by Bram and Ludvigson (1998), there is shown the precise wording of each of the five questions that feature in the University of Michigan's survey.
[41] For the purpose of clarification, a downward movement in the ICS would signify an increased probability of financial stress or uncertainty, and so a reduced willingness to be committed to holding a good that could not be readily converted into a monetary form.

consumption at different levels of aggregation.[42] Huth, Eppright and Taube (1994) also adopted a VAR approach for the purpose of analysis, in conjunction with monthly US data, yet elected to examine the explanatory capability of the ICS alongside that of the Conference Board's consumer confidence indicator and the expectations components of the respective indices. Evidence was obtained of Granger-causality extending from the ICS to sales of both automobiles and durable goods, but not those of non-durable goods. Consequently, should these results be viewed as being representative, they would imply that the greater (smaller) is the contribution of spending on durable goods towards the overall consumption by the personal sector, the more (less) relevant is the ICS (and, by extension, the GfK sentiment measure) for forecasting.

Finally, attention turns to the Netherlands, for which, over the forecast interval as a whole, consumption fell by 1.28 per cent. This overall decline is attributable to the behaviour of household expenditure over the period, 2008q4 – 2012q4, during which ten out of the sixteen quarterly movements were downwards. In contrast, from 2012q4 to 2016q1, spending generally grew, with eleven out of thirteen changes being upwards. Consequently, it would seem that, for the Netherlands, sufficient variability was exhibited in the data to be able to distinguish between the forecast accuracy of the various models. Table 2D shows three significant values of the $MSE - t$ statistic, each one of which involves (in a positive way) the GfK index. Also, Table 3 indicates that both the AR process and the equation including CCI* are omitted from the 90 per cent MCS. Furthermore, the functions incorporating the CCI and Q2 are not retained when constructing the 75 per cent set.

Still, with reference to the Netherlands, the GfK index is associated with a much lower RMSPE than either the CCI or CCI*. Contributing towards this result is possibly the

[42] Corresponding to three levels of aggregation, this study featured five categories of personal spending: consumer durable goods; automobiles; furniture and household furnishings; appliances; refrigerators. Forecasts were generally better for the highest level of aggregation.

rather sluggish behaviour that is displayed by consumption throughout the forecast interval. As has been mentioned above, only towards the end of the period does there occur a sustained increase in expenditure, which still does not completely offset the earlier fall. It may be recalled that the values of the CCI and CCI* are founded upon solely answers to forward-looking questions, which, in this context, possibly constitutes a deficiency, respecting that there are reasons for believing that the layperson will be prospectively overoptimistic about future macroeconomic and microeconomic circumstances. In particular, Bovi (2009) has maintained that irrational exuberance can stem from a combination of two factors: an overconfidence that identified favourable patterns in the past will be repeated in the future; and an illusion that he/she is able to exert an element of control over his/her own future situation. Recall that, in contrast to the CCI and CCI*, values of the GfK index are based additionally upon backward reflection, which would seem to offer the potential for any unwarranted buoyancy to be partly counteracted.

V. Concluding Remarks

The broad objective of this article has been to progress research in the area of the relationship between household consumption expenditure and consumer/economic sentiment. Within earlier empirical studies, the forecasting performances of a variety of different confidence indicators have been explored. Consequently, the more specific aim of this paper has been to undertake a formal comparison of the predictive accuracy of eight regression models, including seven distinct gauges of sentiment, with a particular interest in whether or not the EU's CCI merits its status as the headline measure. The results which have been reported have followed from the application of MSE-t tests and the MCS procedure. Implementation of the latter was regarded as especially apt in an environment of several competing models. Time-series data were collected and analysed on France, Germany, Italy, the Netherlands and

the UK. A multi-country investigation was favoured in order to reduce the risk of producing purely insignificant findings on account of an absence of information content in the data.

The results that were obtained on the five countries were far from uniform. To be more precise, the lack of variability in the growth of consumption over the forecast interval in Germany and the UK seemed to prevent the statistical methods from being able to discriminate between the regression functions. In France, the close relationship between the movements in the confidence indicators over the prediction period also limited the potential to draw a distinction between seven of the specifications. In contrast, though, the fluctuations that were exhibited by household expenditure in Italy and the Netherlands over the interval, 2008q4 – 2016q1, presented a greater opportunity to detect differences in predictive performance. Indeed, in Italy, the highly visible cyclical behaviour of consumer spending helped to produce a situation in which there was only one surviving specification in the 75 per cent MCS.

From the pairwise tests which were conducted, none of the values of $MSE - t$, which were observed to be significant at the five per cent level, were found to involve adversely the CCI. Also, the EU's favoured measure of consumer confidence entered every country's 90 per cent MCS. While this statistical evidence can be used as justification for the continued use of the CCI as the headline indicator, the results have shown that there have been occasions on which this has been outperformed by its rivals. For example, although the CCI contributes towards the most accurate predictions of the growth of consumption in Italy and the UK, the GfK index occupies a pre-eminent position in the Netherlands.

In Italy, the small proportion of total household expenditure that has been devoted to durable goods has had the effect of disadvantaging the GfK indicator, compared to the CCI, given its part dependence on responses to Question 8 in the EU's consumer survey. In contrast, in the Netherlands, the rather sluggish behaviour that was exhibited by consumer

spending over the forecast interval acted as a hindrance to an aggregate measure of sentiment, the value of which was founded upon solely answers to forward-looking questions.

The general conclusion that can be reached in this study is that the statistical evidence is insufficiently strong to be able to recommend a discontinuation of the use of the CCI as the EU's headline indicator of consumer confidence. However, there are reasons for believing that, at some point in the future, revisions may be advisable to the manner of its construction. First, it appears to be a feature of developed nations that, over the course of time, in conjunction with a trend increase in real household expenditure, a greater percentage of total spending by the personal sector is devoted to durable goods. For example, in the UK, this figure has risen from 3.9 per cent in 1985q1, to 5.6 per cent in 2000q1, to 9.5 per cent in 2016q1. Ultimately, then, a dependence upon the responses to Question 8 in the EU's consumer survey would represent a virtue, rather than an impediment towards achieving accurate forecasts. Second, a fundamental objective of governments and monetary authorities is macroeconomic stability. Should eventually success be achieved in terms of eliminating boom and bust then a more useful measure of sentiment may be an indicator that relies upon answers to backward- as well as forward-looking questions.

Acknowledgements

The authors would like to express their thanks to Robin Lumsdaine for providing comments on an earlier draft of the paper and Sylvain Barde for discussions concerning the MCS procedure.

Disclosure Statement

No potential conflict of interest was reported by the authors.

References

Acemoglu, Darren, and Andrew Scott. 1994. "Consumer Confidence and Rational Expectations: Are Agents' Beliefs Consistent with the Theory?" *Economic Journal* 104 (422): 1-19.

Al-Eyd, Ali, Ray Barrell, and E. Philip Davis. 2009. "Consumer Confidence Indices and Short-Term Forecasting of Consumption." *The Manchester School* 77 (1): 96-111.

Barsky, Robert B., and Eric R. Sims. 2012. "Animal Spirits and the Meaning of Innovations in Consumer Confidence." *American Economic Review* 102 (4): 1343-1377.

Boone, Laurence, Nathalie Girouard, and Isabelle Wanner. 2001. "Financial Market Liberalisation, Wealth and Consumption." *OECD Economics Department Working Papers*, No. 38, Paris: OECD Publishing.

Bovi, Maurizio. 2009. "Economic versus Psychological Forecasting. Evidence from Consumer Confidence Surveys." *Journal of Economic Psychology* 30 (4): 563-574.

Bram, Jason, and Sydney Ludvigson. 1998. "Does Consumer Confidence Forecast Household Expenditure? A Sentiment Index Horse Race." *Economic Policy Review* 4: 59-78.

Carroll, Christopher D., Jeffrey C. Fuhrer, and David W. Wilcox. 1994. "Does Consumer Sentiment Forecast Household Spending? If So, Why?" *American Economic Review* 84 (5): 1397-1408.

Clark, Todd, and Michael McCracken. 2013. "Advances in Forecast Evaluation." In Elliott, Graham, and Allan Timmermann (Eds.), *Handbook of Economic Forecasting*, Vol. 2B, 1107-1201, Elsevier.

Clark, Todd E., and Michael W. McCracken. 2002. "Evaluating Long-Horizon Forecasts." Federal Reserve Bank of Kansas City, Working Paper No. 01-14.

Cotsomitis, John A., and Andy C.C. Kwan. 2006. “Can Consumer Confidence Forecast Household Spending? Evidence from the European Commission Business and Consumer Surveys.” *Southern Economic Journal* 72 (3): 597-610.

Diebold, Francis X., and Roberto S. Mariano. 1995. “Comparing Predictive Accuracy.” *Journal of Business and Economic Statistics* 13 (3): 253-263.

Easaw, Joshy Z., Dean Garratt, and Saeed M. Heravi. 2005. “Does Consumer Sentiment Accurately Forecast UK Household Consumption? Are There Any Comparisons to be Made with the US?” *Journal of Macroeconomics* 27 (3): 517-532.

Easaw, Joshy Z., and Saeed M. Heravi. 2004. “Evaluating Consumer Sentiments as Predictors of UK Household Consumption Behavior. Are They Accurate and Useful?” *International Journal of Forecasting* 20 (4): 671-681.

Gausden, Robert, and Mohammad S. Hasan. 2016. “Would Information on Consumer Confidence Have Helped to Predict UK Household Expenditure during the Recent Economic Crisis?” *Applied Economics* 48 (18): 1695-1709.

Gausden, Robert, and Mohammad S. Hasan. 2012. “A Comparison of Consumer and Retail Trade Confidence Indicators for Predicting Household Expenditure in the UK.” *The Empirical Economics Letters* 11 (7): 669-676.

Giacomini, Raffaella, and Halbert White. 2006. “Tests of Conditional Predictive Ability.” *Econometrica* 74 (6): 1545-1578.

Hansen, Peter R., Asger Lunde, and James M. Nason. 2011. “The Model Confidence Set.” *Econometrica* 79 (2): 453-497.

Harvey, David, Stephen Leybourne, and Paul Newbold. 1997. “Testing the Equality of Prediction Mean Squared Errors.” *International Journal of Forecasting* 13 (2): 281-291.

Huth, William L., David R. Eppright, and Paul M. Taube. 1994. “The Indexes of Consumer Sentiment and Confidence: Leading or Misleading Guides to Future Buyer Behavior.” *Journal of Business Research* 29 (3): 199-206.

Kumar, V., Robert P. Leone, and John N. Gaskins. 1995. “Aggregate and Disaggregate Sector Forecasting Using Consumer Confidence Measures.” *International Journal of Forecasting* 11 (3): 361-377.

Lahiri, Kajal, George Monokroussos, and Yongchen Zhao. 2016. “Forecasting Consumption: The Role of Consumer Confidence in Real Time with Many Predictors.” *Journal of Applied Econometrics* 31 (7): 1254-1275.

Mishkin, Frederic S., Robert Hall, John Shoven, Thomas Juster, and Michael Lovell. 1978. “Consumer Sentiment and Spending on Durable Goods.” *Brookings Papers on Economic Activity*: 1978 (1): 217-232.

Nahuis Niek J., and W. Jos Jansen. 2004. “Which Survey Indicators are Useful for Monitoring Consumption? Evidence from European Countries.” *Journal of Forecasting* 23 (2): 89-98.

Starr, Martha A. 2012. “Consumption, Sentiment and Economic News.” *Economic Inquiry* 50 (4): 1097-1111.

West, Kenneth D. 1996. “Asymptotic Inference about Predictive Ability.” *Econometrica* 64 (5): 1067-1084.

Table 1. Descriptive Statistics for Measures of Confidence

	France		Germany		Italy		Netherlands		UK	
Variable	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
CCI	-18.99	8.0111	-7.273	8.9880	-15.39	8.4538	1.8485	12.351	-8.800	8.6482
CCI*	-18.67	10.144	-11.91	11.089	-14.00	9.2489	-8.019	15.578	-11.13	9.8766
GfK	-18.26	8.4313	-7.954	10.869	-19.69	10.527	-4.959	14.844	-9.121	10.999
Q2	-3.700	4.8873	-2.296	4.8586	-2.573	5.2998	0.6149	6.7300	0.9501	7.9028
Q4	-20.11	10.464	-11.20	12.974	-9.426	11.078	-6.891	17.074	-11.06	11.248
Q8	-15.49	11.678	-3.220	12.609	-31.99	16.901	-0.795	21.507	2.4851	15.927
ESI	100.53	9.8620	100.81	9.4544	100.73	9.784	100.16	9.231	101.66	10.438

Note: This table presents the values of the sample mean and standard deviation (S.D.) for each of the seven confidence variables and five countries. The common sample period is 1985q1-2016q1.

Table 2A. Computed Values of the MSE-t Statistic in relation to France

Model	AR	CCI	CCI*	GfK	Q2	Q4	Q8	ESI
AR	-							
CCI	2.3686*	-						
CCI*	2.1044*	1.2856	-					
GfK	0.6426	-0.3796	-0.9177	-				
Q2	0.0752	-0.7502	-1.1066	-0.5720	-			
Q4	1.6612	1.0269	0.6492	1.4634	1.4905	-		
Q8	1.5044	0.6287	-0.0161	1.1765	1.0539	-0.5261	-	
ESI	0.2164	-0.9378	-1.1108	-0.4196	0.0748	-1.1527	-1.1576	-

Notes: Each cell contains the computed value of the MSE-t statistic.
A positive (negative) value is an indication that the model which features in the top row has a higher (lower) MSE than the model which enters the first column.
For a two-tailed test, the relevant critical values are: 1.96 (5 per cent level of significance); 2.575 (1 per cent).
* denotes a significant value at the 5 per cent level; + denotes a significant value at the one per cent level.

Table 2B. Computed Values of the MSE-t Statistic in relation to Germany

Model	AR	CCI	CCI*	GfK	Q2	Q4	Q8	ESI
AR	-							
CCI	-0.2011	-						
CCI*	-0.3033	-1.7956	-					
GfK	0.1238	0.5694	0.7357	-				
Q2	-0.1218	0.1038	0.2104	-0.2433	-			
Q4	-0.4878	-1.0896	-0.8228	-1.2240	-0.4198	-		
Q8	0.3396	0.2866	0.3893	-0.0424	0.2822	0.5639	-	
ESI	0.2490	0.7891	0.9524	0.2960	0.2939	1.4975	0.1773	-

For notes, see Table 2A.

Table 2C. Computed Values of the MSE-t Statistic in relation to Italy

Model	AR	CCI	CCI*	GfK	Q2	Q4	Q8	ESI
AR	-							
CCI	2.1391*	-						
CCI*	1.8233	-1.9002	-					
GfK	-1.1562	-2.8057+	-2.6499+	-				
Q2	0.8949	-1.7320	-0.5790	2.1312*	-			
Q4	-0.5062	-2.3565*	-2.0639*	1.0532	-1.7554	-		
Q8	-2.6537+	-2.9774+	-3.0250+	-2.5532*	-2.8206+	-2.7229+	-	
ESI	-1.6362	-2.2557*	-2.1650*	-1.0433	-1.8874	-1.5162	2.9687+	-

For notes, see Table 2A.

Table 2D. Computed Values of the MSE-t Statistic in relation to the Netherlands

Model	AR	CCI	CCI*	GfK	Q2	Q4	Q8	ESI
AR	-							
CCI	0.8953	-						
CCI*	0.6500	-1.2213	-					
GfK	2.4456*	1.9262	2.2681*	-				
Q2	0.3205	-0.9111	-0.4439	-2.2849*	-			
Q4	0.6676	-0.1101	0.3357	-1.7083	0.5722	-		
Q8	0.3801	-0.2869	-0.1172	-1.3510	0.0800	-0.2371	-	
ESI	1.3919	1.0326	1.5602	-0.4253	1.2331	1.0354	0.6645	-

For notes, see Table 2A.

Table 2E. Computed Values of the MSE-t Statistic in relation to the UK

Model	AR	CCI	CCI*	GfK	Q2	Q4	Q8	ESI
AR	-							
CCI	1.6322	-						
CCI*	1.3952	-0.4336	-					
GfK	1.1636	-1.2059	-0.9561	-				
Q2	1.7387	-0.1340	0.1924	0.9388	-			
Q4	0.0775	-2.2275*	-1.7870	-1.2686	-1.9709*	-		
Q8	0.3497	-1.2994	-1.0267	-0.9423	-1.3756	0.2842	-	
ESI	0.2049	-1.6469	-1.4342	-0.9297	-1.4275	0.1339	-0.1485	-

For notes, see Table 2A.

Table 3. MCS for One-Step-Ahead Forecasts of $\Delta \ln(Cons_t)$

	Country									
	France		Germany		Italy		Netherlands		UK	
Model	RMSPE x 10^2	prob. value	RMSPE x 10^2	prob. Value	RMSPE x 10^2	prob. value	RMSPE x 10^2	prob. value	RMSPE x 10^2	prob. value
AR	0.52	(0.07)	0.47	(0.99)**	0.51	(0.18)*	0.69	(0.08)	0.77	(0.35)**
CCI	0.49	(0.53)**	0.48	(0.95)**	0.45	(1.00)**	0.66	(0.15)*	0.61	(1.00)**
CCI*	0.47	(0.73)**	0.49	(0.33)**	0.47	(0.18)*	0.67	(0.08)	0.63	(0.86)**
GfK	0.50	(0.53)**	0.47	(0.99)**	0.55	(0.10)	0.60	(1.00)**	0.67	(0.52)**
Q2	0.52	(0.53)**	0.48	(0.99)**	0.49	(0.18)*	0.68	(0.12)*	0.62	(0.88)**
Q4	0.46	(1.00)**	0.50	(0.71)**	0.53	(0.14)*	0.66	(0.29)**	0.76	(0.18)*
Q8	0.47	(0.73)**	0.47	(0.99)**	0.67	(0.08)	0.68	(0.30)**	0.73	(0.43)**
ESI	0.51	(0.53)**	0.46	(1.00)**	0.59	(0.18)*	0.62	(0.68)**	0.75	(0.36)**

Notes: RMSPE represents the Root Mean Square Prediction Error.
* signifies that the model enters the 90 per cent MCS. ** signifies that the model enters the 75 per cent MCS.

Appendix

Table A1. Details of Data on Household Consumption Expenditure

Country	*Consumption Variable*	*Datastream Codename*	*Availability*
United Kingdom	Final Consumption Expenditure, Households and Non-profit Institutions Serving Households (Millions, pounds, 2013 prices)	UKCNPER.D	1985q1-2016q1
France	Final Consumption Expenditure, Households and Non-profit Institutions Serving Households (Millions, Euros, 2010 prices)	FRCNPER.D	1985q1-2016q1
Italy	Final Consumption Expenditure, Households and Non-profit Institutions Serving Households (Millions, Euros, 2010 prices)	ITCNPER.D	1996q1-2016q1
	Final Consumption Expenditure, Households (Millions, Euros, 2005 prices)	ITCNPERND	1991q1-2014q2
	Final Consumption Expenditure, Households (Millions, Euros, 2000 prices)	ITCNPEG0D	1985q1-2011q3
Germany	Final Consumption Expenditure, Households and Non-profit Institutions Serving Households (Billions, Euros, 2010 prices)	BDCNPER.D	1991q1-2016q1
Netherlands	Final Consumption Expenditure, Households and Non-profit Institutions Serving Households (Millions, Euros, 2010 prices)	NLCNPER.D	1996q1-2016q1
	Final Consumption Expenditure, Households and Non-profit Institutions Serving Households (Millions, Euros, 2005 prices)	NLCNPERND	1988q1-2014q1

Table A2. Questions Relating to the Joint Harmonised EU Consumer Survey

Question 1	How has the financial situation of your household changed over the last 12 months?
Question 2	How do you expect the financial position of your household to change over the next 12 months?
Question 3	How do you think the general economic situation in the country has changed over the past 12 months?
Question 4	How do you expect the general economic situation in this country to develop over the next 12 months?
Question 5	How do you think that consumer prices have developed over the last 12 months?
Question 6	By comparison with the past 12 months, how do you expect that consumer prices will develop in the next 12 months?
Question 7	How do you expect the number of people unemployed in this country to change over the next 12 months?
Question 8	In view of the general economic situation, do you think that now is the right moment for people to make major purchases such as furniture, electrical/electronic devices, etc.?
Question 9	Compared to the past 12 months, do you expect to spend more or less money on major purchases (furniture, electrical/electronic devices etc.) over the next 12 months?
Question 10	In view of the general economic situation, do you think that now is a very good/fairly good/not a good/a very bad moment to save?
Question 11	Over the next 12 months, how likely is it that you save any money?
Question 12	Which of these statements best describes the current financial situation of your household? - We are saving a lot - We are saving a little - We are just managing to make ends meet on our income - We are having to draw on our savings - We are running into debt

Source: European Commission Directorate-General for Economic and Financial Affairs, The Joint Harmonised EU Programme of Business and Consumer Surveys, User Guide (updated February 2017). (https://ec.europa.eu/info/sites/info/files/bcs_user_guide_en_0.pdf)