Received: 4 April 2021 | Revised: 23 September 2021 | Accepted: 23 November 2021

DOI: 10.1111/manc.12395

ORIGINAL ARTICLE

The Manchester School WILEY

A reappraisal of Katona's adaptive theory of consumer behaviour using U.K. data

Robert Gausden[1] | **Mohammad Hasan**[2]

[1]Subject Group of Economics and Finance, University of Portsmouth, Portsmouth, UK

[2]Kent Business School, University of Kent, Canterbury, UK

Correspondence
Robert Gausden, Subject Group of Economics and Finance, University of Portsmouth, Portsmouth PO1 3DE, UK.
Email: Robert.Gausden@port.ac.uk

Abstract

The objective of this paper is to conduct a reappraisal of Katona's (1968) adaptive theory of consumer behaviour, which maintains that discretionary consumption is partly determined by attitudes and expectations of households. Initially, using UK data, we follow Katona by empirically examining whether changes in personal expenditure on durable goods are connected to earlier movements in consumer confidence. Evidence of a lack of a stable relationship between these two variables encourages us to perform a disaggregated analysis involving 111 components of four different forms of consumption, which enables construction of an aggregate measure of discretionary spending. We find that sufficient criteria are satisfied for the sentiment index to be accepted as a reliable predictor of the growth of gratuitous expenditure. In conclusion, then, it would seem that the validity of Katona's theory can be revived if we are prepared to discard the assumption that durable goods' consumption is synonymous with discretionary spending.

KEYWORDS

consumer confidence, discretionary expenditure, durable goods, household consumption, Katona's adaptive theory

JEL CLASSIFICATION

E21; E27; E71

1 | INTRODUCTION

In a seminal journal article, Katona (1968) presented an adaptive theory of consumer behaviour, which he intended to rival traditional theories, such as the Permanent Income Hypothesis. A distinguishing feature of Katona's theory is the contention that consumers' discretionary expenditures are governed by both the ability and the willingness to buy. The ability to buy is largely represented by income, financial assets and the availability of credit. In contrast, the willingness to buy is mainly decided by attitudes and expectations concerning personal finances and the general economic situation.

In related empirical work, Katona (1967, 1968) employed as a summary measure of attitudes and expectations the University of Michigan's Index of Consumer Sentiment (ICS). Also, he elected to proxy discretionary household spending by expenditure on durable goods. Conducting regression analysis, using US data from 1952 to 1966, Katona (1967) found evidence of a statistically significant effect of each of disposable personal income and the ICS on the consumption of durable goods. In addition, he was able to point to an impressive performance of the ICS in forecasting turning points. However, there have occurred several subsequent econometric studies which have shown the influence of a measure of consumer confidence on durable goods expenditure to be dependent upon the context in which the analysis takes place. These include the investigations by Bram and Ludvigson (1998), Throop (1992), Baghestani and Kherfi (2015), Ahmed and Cassou (2016), Gausden and Hasan (2016), and Easaw and Heravi (2004). On this basis, consumer sentiment cannot be regarded as a wholly reliable predictor of durable goods' consumption, which, in turn, would seem to cast doubt over the validity of Katona's hypothesis.

This failure to find a consistent relationship between household spending on durable goods and consumer confidence motivates us, in this paper, to provide a reappraisal of Katona's theory. In particular, our angle of approach is to question whether the consumption of durable goods serves as a suitable proxy for discretionary household expenditure. Indeed, Katona (1968), himself, admitted that, amongst the purchases of durable goods, there may be some items which can be classed as essential. At the same time, then, it would seem to be reasonable to suppose that there are elements of spending on other types of goods and services which are viewed as optional. The fundamental contribution which is made by our paper is to undertake a formal statistical analysis involving 111 different types of personal consumption in the UK to determine which of these can be categorised as discretionary. It transpires that 45 components of consumption qualify as being non-obligatory, comprising just under 33 per cent of UK household expenditure in 2015. Interestingly, out of all of the gratuitous spending that was undertaken by households in that year, less than 15 per cent was devoted to durable goods. Also, just over half of all of durable goods' consumption merited the description of discretionary.

Consequently, for the purpose of performing what we consider to be a fairer assessment of Katona's theory, we create an aggregate measure of discretionary household expenditure (DISC). The framework that is chosen in order to investigate the relationship between the latter and our preferred consumer confidence indicator is the well-reputed autoregressive distributed-lag (ARDL) model that was constructed by Bram and Ludvigson (1998), which is estimated using UK quarterly data. In order for the sentiment variable to be accepted as a reliable predictor of DISC, three conditions need to be satisfied. First, following estimation over the full sample period, the collective effect of the past values of the consumer confidence variable on the growth of DISC should be statistically significant. Second, this statistical significance should be largely preserved, having conducted estimation over suitable sub-periods, and there should be a lack of evidence of temporal instability.

Finally, the accuracy of post-sample forecasts of changes in DISC ought to benefit appreciably from the inclusion of a representation of consumer sentiment in the respective regression equation.

From our empirical analysis, we find that, when DISC is employed as the consumption variable, the above three conditions are indeed satisfied. In contrast, with household expenditure on durable goods replacing DISC, only the first of the requirements is met. Thus, it follows that, were spending on durable goods to have been selected as a proxy for discretionary consumption, Katona's adaptive theory would have been falsely rejected. Furthermore, for the purpose of guiding macroeconomic policy, the importance of shifts in consumer confidence would be seriously underestimated. For one reason, changes in attitudes and/or expectations would not always be trusted to provide an indication of future movements in household expenditure. Also, even if shifts in sentiment were regarded as being informative, it must be respected that the consumption of durable goods amounted to less than 10 per cent of overall spending by the UK personal sector in 2015, whereas, for DISC, the corresponding figure was nearly 33 per cent.

The paper proceeds in the following manner. Section 2 discusses the literature which provides a background to this study. Section 3 details the empirical methodology that is favoured, while exploring the usefulness of a gauge of consumer confidence in accounting for the growth of household expenditure on durable goods. Section 4 is devoted to a disaggregated analysis for the purpose of determining which of 111 components of personal spending can be classed as inessential, thereby giving rise to the construction of a variable to represent total discretionary expenditure. In Section 5, the empirical methodology is applied to this composite measure of consumption in order to establish whether consumer sentiment is better suited to forecasting changes in discretionary spending. Finally, section 6 contains a summary of the principal findings and some concluding comments.

2 | LITERATURE REVIEW

In the second part of this article, we supply a brief review of the literature which provides a background to the econometric analysis that is undertaken in later sections. The emphasis is deliberately upon those empirical studies which have investigated the relationship between consumer confidence and durable goods' consumption, given our contention that the latter may be an inadequate proxy for discretionary household expenditure.

An appropriate starting point would seem to be the paper by Bram and Ludvigson (1998), granted that the ARDL model which they constructed is chosen as the framework for conducting our own applied research. One of the empirical questions that Bram and Ludvigson (1998) sought to answer was whether or not an indicator of consumer sentiment contained information about future household spending in the US which was supplementary to that which was incorporated within some established macroeconomic and financial variables. The majority of US empirical studies in this area have favoured the use of the University of Michigan's ICS (or a component of this) to convey the attitudes and expectations of households. In contrast, the paper by Bram and Ludvigson (1998) involved not only the ICS but also the Conference Board's consumer confidence index (CBI).[1] Their ARDL models featured five different consumption variables, which included spending on each of motor vehicles and all other durable goods, but not durable goods in total.

[1]In the second section of their paper, Bram and Ludvigson (1998) indicate the key differences between the ICS and the CBI.

Bram and Ludvigson (1998) undertook both a within- and a post-sample analysis using US quarterly data from 1968Q1 to 1996Q3. In general, the CBI was found to have greater explanatory or predictive power than the ICS. However, even the performance of the CBI was seen to be somewhat variable. With reference to the within-sample analysis, for both categories of durable goods, the addition of the CBI to the respective baseline equation was able to achieve a clear improvement in the fit of the data on the dependent variable. Concerning the post-sample analysis, for motor vehicles expenditure, the inclusion of the CBI in the ARDL model repeatedly increased the accuracy of one-quarter-ahead forecasts, although the benefit from this extension was much more apparent over the sub-period 1982Q1–1989Q4 than 1990Q1–1996Q3. In contrast, for all other durable goods, the same augmentation to the distributed-lag equation produced a very slight gain during the 1990s, yet led to much poorer quality predictions during the 1980s.

In terms of the ability of a measure of consumer confidence to forecast household spending on durable goods, the somewhat mixed results that were obtained by Bram and Ludvigson (1998) have also been a feature of other US econometric studies. For instance, Throop (1992) established that a contribution from the ICS was only advantageous when unusual events were occurring,[2] while Baghestani and Kherfi (2015) found that durable consumption reacted merely to falls but not rises in consumer sentiment. Furthermore, Ahmed and Cassou (2016), adopting the framework of a threshold local projection model, observed a stronger response of durable goods expenditure to an innovation in the ICS during an economic expansion than a contraction.

The same, rather inconsistent findings have also emerged from analyses of UK data that were performed by Gausden and Hasan (2016) and Easaw and Heravi (2004). In the paper by Gausden and Hasan (2016), the favoured measure of consumer sentiment was the European Commission's Consumer Confidence Indicator (CCI).[3] While it was possible to present evidence to show how the involvement of the latter improved out-of-sample forecasts of the growth of spending on durable goods, this only applied to a period which included the financial crisis (2008Q1–2013Q1) and did not have extension to an earlier, less turbulent phase (2002Q4–2007Q4). In contrast, Easaw and Heravi (2004) elected to incorporate in their regression equations the GfK sentiment index. For five categories of household expenditure, one-quarter-ahead predictions were generated recursively over two distinct intervals, 1990Q1–1994Q4 and 1995Q1–2000Q4, which were characterised by adverse and rising consumer confidence, respectively. Results indicated that the presence of the GfK index served to enhance significantly the accuracy of the forecasts of the growth of spending on all durable goods, motor vehicles, and other durable goods. However, the qualification must be added that the observed improvement was limited to merely the later of the two time periods.

Finally, Easaw et al. (2005) also conducted an evaluation of how well consumer sentiment predicts household consumption expenditure in the UK. Their study involved three consumption variables and two measures of consumer confidence. An initial model related the change in the logarithm of consumption to past values of itself, the growth of labour income, and a consumer sentiment variable. Following estimation over the period, 1982Q1–1999Q4, the lags on consumer confidence were found to be statistically significant when spending on durable goods served as the consumption variable and either the GfK or MORI indices fulfilled the role of the sentiment measure. It should be noted, though, that this study did not attempt a post-sample analysis.

[2]To provide a precise example, Throop (1992) discovered that more accurate predictions could be achieved over the interval, 1990Q3–1991Q3, by virtue of involving the ICS in the analysis. This period was notorious for incorporating the invasion of Kuwait by Iraq and the subsequent military intervention by the US and its allies.

[3]In Section 3 of our paper, we highlight the differences in the construction of three available measures of consumer confidence for the UK: the GfK index; the CCI; and the MORI index.

3 | DURABLE GOODS' CONSUMPTION AND CONSUMER CONFIDENCE

In this section, we explore the predictive capability of a measure of consumer confidence with respect to the growth of UK household consumption expenditure on durable goods in order to see how our results compare with those obtained in earlier studies. Initially, a within-sample analysis is undertaken. More specifically, in accordance with the approach that has been adopted by Bram and Ludvigson (1998), Carroll et al. (1994) and Easaw and Heravi (2004), a baseline ARDL model is first constructed, which does not incorporate an attitudinal variable. Subsequently, this equation is extended by including past values of an indicator of consumer sentiment. The usefulness of the consumer confidence variable can then be ascertained by comparing the values of an adjusted R-squared statistic. Also, in conjunction with the augmented regression equation, there is the facility to perform an F test of the joint null hypothesis that past changes in sentiment have no effect on the dependent variable.

The regression function which fulfils the role of the baseline model of durable goods' consumption is shown below as Equation (1).

$$\Delta\log\left(DUR_t\right) = B_0 + \sum_{i=1}^{n} B_{1i}\Delta\log\left(DUR_{t-i}\right) + \sum_{i=1}^{n} B_{2i}\Delta\log\left(RHDI_{t-i}\right) + \sum_{i=1}^{n} B_{3i}\Delta\log\left(FTSE_{t-i}\right) + \sum_{i=1}^{n} B_{4i}\Delta TB_{t-i} + \varepsilon_t \tag{1}$$

With reference to this equation: DUR = UK household consumption expenditure on durable goods (seasonally adjusted, constant (2015) prices); $RHDI$ = UK real household disposable income (seasonally adjusted); $FTSE$ = the FTSE All Share Price Index, expressed in real terms; TB = UK three-month Treasury bill rate of discount.[4]

The extended regression model is indicated by Equation (2), below, which features additionally past values of the (seasonally-adjusted) GfK index of consumer sentiment.

$$\Delta\log\left(DUR_t\right) = B_0 + \sum_{i=1}^{n} B_{1i}\Delta\log\left(DUR_{t-i}\right) + \sum_{i=1}^{n} B_{2i}\Delta\log\left(RHDI_{t-i}\right) + \sum_{i=1}^{n} B_{3i}\Delta\log\left(FTSE_{t-i}\right) + \sum_{i=1}^{n} B_{4i}\Delta TB_{t-i} + \sum_{i=1}^{n} B_{5i} GfK_{t-i} + \varepsilon_t \tag{2}$$

It should be noted that the variables enter Equations (1) and (2) in such a form that the associated time series are stationary. Also, in choosing the number of quarterly lags to admit on the variables, the default selection is $n = 4$. However, in accordance with the approach of Bram and Ludvigson (1998), a suitable comparison is made of values of information criteria to confirm that longer lags are unnecessary.

It is evident, then, that the GfK index is our preferred summary measure of households' attitudes and expectations. The value of this indicator is based upon responses to the following five survey questions.

[4]Within the Appendix to this paper, information is provided on the construction of these variables and the related data sources.

Current conditions

1. How does the financial situation of your household now compare with what it was 12 months ago?
2. How do you think the general economic situation has changed over the last 12 months?
3. Do you think there are benefits in people making major purchases, such as furniture, washing machines, TV sets, at the present time?

Expected future conditions

4. How do you think the financial position of your household will change over the next 12 months?
5. How do you think the general economic situation will develop over the next 12 months?

With regard to questions 1, 2, 4 and 5, the potential answers consist of: a lot better (PP); a little better (P); the same (N); a little worse (M); and a lot worse (MM). In contrast, for question 3, there are only three possible responses: yes, it is the right moment now (PP); it is neither the right nor the wrong moment (N); no, it is not the right moment now (MM).[4] Numerical values are allocated to the five categories of answer, which consist of: PP = 1; P = 0.5; N = 0; M = −0.5; and MM = −1. For each question, a net balance can be achieved by adding together the scores of all of the participants in the survey and expressing the result as a percentage of the total number of responses. A simple arithmetic average of the five net balances yields the value of the GfK index, which is compelled to have the bounds of −100 and 100, by virtue of the manner of its construction.

The reason for favouring the GfK index over, for example, the CCI and the MORI indicator is its part reliance upon a question that seems to be particularly apt for the purpose of explaining household expenditure on durable goods. This is, of course, question 3, above, which makes direct reference to major purchases, citing three types of durable goods as examples. The CCI is similar to the GfK index in terms of the way in which it is assembled. The key distinction is that the value of the CCI is founded upon the responses to four forward-looking questions, which are concerned with developments over the next twelve months to the individual's financial position, the general economic situation, unemployment, and the potential to make savings. In contrast, Ipsos MORI produce an Economic Optimism Index, which is simply based on the answers to the single question, "Do you think that the general economic condition of the country will improve, stay the same or get worse over the next twelve months?"

Ordinary Least Squares estimation is applied to both Equations (1) and (2) using quarterly data which extend from 1986Q2 to 2016Q3.[5] The results of the estimation and subsequent tests are shown in the first two columns of Table 1. Upon studying the latter, it is apparent that the extension of the baseline equation to include past values of the GfK index serves to increase the value of the adjusted R-squared statistic by 0.0581. Also, within Equation (2), out of the five potential determinants of $\Delta\log(DUR_t)$, only the lags on the confidence measure are exerting a statistically significant effect.

[4]Additionally, for any of the questions, there is the capacity to answer "don't know".

[5]The start date is restricted by quarterly data being available on UK household consumption expenditure on durable goods from 1985Q1. Also, 2016Q3 represents the most recent quarter for which consumption data could be obtained, when the first draft of this paper was being prepared.

TABLE 1 Results obtained from estimating equations for $\Delta\log(DUR_t)$ over the full sample period

Right-hand-side variables	Equation (1)	Equation (2)	Augmented Equation (1)	Augmented Equation (2)
$\Delta\log\left(DUR_{t-i}\right)$	0.2404 (0.3081)	−0.1399 (0.3906)	0.3789 (0.0034)	−0.0077 (0.0117)
$\Delta\log\left(RHDI_{t-i}\right)$	0.2125 (0.3580)	−0.1988 (0.4163)	0.1271 (0.3113)	−0.3562 (0.3043)
$\Delta\log\left(FTSE_{t-i}\right)$	0.0946 (0.1075)	0.0547 (0.3112)	0.1720 (0.0474)	0.1370 (0.1185)
ΔTB_{t-i}	−0.0072 (0.6977)	−0.0073 (0.6207)	−0.0014 (0.8986)	−0.0008 (0.5839)
GfK_{t-i}		0.0009 (0.0428)		0.0010 (0.0104)
Adjusted R-squared	0.0260	0.0841	0.1071	0.1954
BG(4)	7.4553 (0.1137)	7.5905 (0.1078)	2.2410 (0.6915)	3.6901 (0.4496)

Note: All statistical results which are shown in this and other tables are produced using *EViews* 10. All of the equations have been estimated by Ordinary Least Squares over the sample period, 1986Q2–2016Q3. Tests of hypotheses are performed with allowance for heteroskedasticity and autocorrelation in the disturbance terms (via selecting the Newey-West covariance procedure). The augmented equations incorporate additionally two dummy variables to represent the redistribution of expenditure which occurred over 2009 and 2010 as a consequence of the implementation of the Labour Government's vehicle scrappage scheme. The first five rows of the table show the sums of the estimated parameters which are attached to the four lags on the respective variable. The figures in parentheses are the probability values corresponding to an *F* test of the null hypothesis that each of the four parameters is equal to zero. BG(4) is the value of a Breusch-Godfrey chi-square statistic which has been computed for the purpose of testing for up to fourth-order autocorrelation in the disturbance terms. Figures in parentheses are probability values.

The final two columns of Table 1 contain the results which follow from the estimation of adapted versions of Equations (1) and (2). To be more specific, the decision was taken to supplement each of the two regression functions through the addition of two dummy variables. The latter were designed with the intention of capturing the impact on durable goods expenditure of the Labour Government's vehicle scrappage scheme, which was introduced in May 2009 in an attempt to provide a stimulus to the UK's automotive sector, which had endured a steep decline in car sales during the financial crisis. The reason for the creation of two dummy variables was that purchases of motor vehicles were noticeably larger than they would otherwise have been in the final three quarters of 2009, while smaller during the second three months of 2010. Consideration of Table 1 reveals that the incorporation of the two dummy variables has the effect of increasing the values of the adjusted R-squared statistic, by 0.0811 in the case of the baseline equation and 0.1113 for the equation featuring the GfK index. Also, it is apparent that the inclusion in the analysis of the two dummy variables has the effect of sharpening the influence of the past values of consumer sentiment on $\Delta\log(DUR_t)$. As evidence, the consequence of admitting the four lags on the confidence variable is to raise the value of the adjusted R-squared statistic by 0.0883. Also, within the final column of Table 1, it can be seen that the F probability value which is associated with the GfK index is only 0.0104.[6]

Mention was made earlier of there being available two other UK sentiment measures, namely, the CCI and the MORI index. Consequently, it seems to be appropriate, at this stage, to examine how sensitive are these within-sample results to the choice of confidence indicator. The Appendix

[6]Table A1, in the Appendix to this paper, supplies more detail, showing, for the least restricted equation, each of the individual parameter estimates, together with an indication of their statistical significance.

to this paper includes the two tables, Tables A2 and A3, which have the same design as Table A1. More precisely, Table A2 relates to an augmented version of Equation (2), which employs the CCI as a replacement for the GfK index, while, in the case of Table A3, the MORI measure operates as the substitute. With regard to Tables A2 and A3, the most pertinent information lies in the cell in the bottom right-hand corner, for this shows whether collectively the past values of the sentiment measure have any connection with the growth of household expenditure on durable goods. The two F probability values are 0.1439 (Table A2) and 0.1654 (Table A3).[7] Since both are above the conventional significance level of 0.05 then the inference is drawn that neither the CCI nor the MORI index contains useful information for predicting changes in durable goods' consumption beyond that which is incorporated in the lagged values of the four control variables. Consequently, had either the CCI or the MORI index been selected as the preferred gauge of sentiment then it seems that, at the earliest possible stage, the consumer confidence indicator would have been dismissed as a relevant determinant of household spending on durable goods.

From the information that has been presented in Table 1, though, it is possible to infer that consumer sentiment, as represented by the GfK index, is an important predictor of the growth of durable goods' consumption. However, in order to be able to regard this sentiment measure as a dependable leading indicator, there is a need for the significance of this variable to be preserved over sub-periods of the full data set. Consequently, the decision is taken to conduct estimation of Equation (2), using four approximately equal-length sub-periods.[8][9] Following estimation, an F test is performed of the null hypothesis that each of the parameters that are attached to $GfK_{t-i},\ i = 1,\ 2,\ 3,\ 4,$ are equal to zero. The computed values of the statistics, together with the associated probability values, are shown in Table 2.

Table 2 indicates, over the course of time, a progressive reduction in the contribution that is being made by the past values of the GfK index towards explaining the variation in $\Delta\log(DUR_t)$. Over the first sub-period, the computed value of the F statistic is found to be significant at the one per cent level. Although the probability value is higher over the second sub-period, it is still observed to fall below 0.10. In contrast, though, for neither of the final two sub-periods is it possible to reject the null hypothesis, $B_{5i} = 0,\ \ i = 1,\ 2,\ 3,\ 4,$ at a conventional level of significance.

A possible interpretation of these results, which would be consistent with the theme of our paper, is that, over the course of time, spending on durable goods has become an increasingly poor proxy for discretionary consumption expenditure, thereby weakening the association that this has with a measure of consumer confidence. When estimation is undertaken of a discrete threshold ARDL model, 2011Q3 is identified as a period when a structural break occurs in terms of the relationship between $\Delta\log(DUR_t)$ and $GfK_{t-i},\ \ i = 1,\ 2,\ 3,\ 4$. Prior to 2011Q3, the sum of the estimates of $B_{5i},\ i = 1,\ 2,\ 3,\ 4$ is 0.0023. Subsequently, though, every one of the estimates shifts downwards, such that their sum is only 0.0003. Following the application of an F test to assess whether or not the changes in the parameter estimates are significant, the computed value of the statistic is $F(4,\ 95) = 5.1962$, with a probability value of 0.0008.

As a more general check on parameter stability, we perform the well-established CUSUM test (Brown et al., 1975), in conjunction with the augmented version of Equation (2). The CUSUM test is founded upon recursive residuals. The test statistic has the form:

[7]For the MORI index, there appears to be a contradiction between the results shown here and in the paper by Easaw et al. (2005). However, the explanation would seem to lie chiefly with the choice of estimation period. It is noticed that the significance of the lags on the MORI indicator becomes reduced when the analysis extends beyond 2012.

[8]Where the dates of the sub-period justify this, the two dummy variables additionally enter Equation (2).

[9]The choice of sub-periods meant that, in each case, the sample size is 30 or 31. Any fewer than 30 observations would seem to constrain the ability to detect significant relationships.

TABLE 2 Results obtained from estimating Equation (2) over selected sub-periods

Sub-period	*F* statistic	Probability value
1986Q2–1993Q4	$F(4, 10) = 7.6374$	0.0043
1994Q1–2001Q2	$F(4, 9) = 3.2374$	0.0662
2001Q3–2008Q4	$F(4, 9) = 1.3568$	0.3221
2009Q1–2016Q3	$F(4, 8) = 0.3650$	0.8272

Note: All of the equations have been estimated by Ordinary Least Squares. Tests of hypotheses are performed with allowance for heteroskedasticity and autocorrelation in the disturbance terms (via selecting the Newey-West covariance procedure). For the final sub-period, 2009Q1–2016Q3, Equation (2) incorporates additionally two dummy variables to represent the redistribution of expenditure which occurred over 2009 and 2010 as a consequence of the implementation of the Labour Government's vehicle scrappage scheme. The value of the *F* statistic corresponds to a test of the joint null hypothesis that each of the parameters which are attached to the lags on the GfK index is equal to zero, i.e., $B_{5i} = 0$, $i = 1, 2, 3, 4$.

$$W_t = \sum_{r=k+1}^{t} \frac{w_r}{s}, \qquad t = k+1, k+2, \cdots\cdots\cdots, T,$$

where w_r denotes a recursive residual and s signifies the standard deviation of the residuals. On the basis of the null hypothesis, which maintains that the parameters of the model are constant, W_t has an expected value of zero. If the computed value of the statistic deviates significantly from zero then the null hypothesis must be rejected.

Figure 1, shows a time plot of W_t, together with a pair of five per cent level of significance lines. Should the computed value of the test statistic stray beyond either of the two significance bands then the inference is drawn that the respective model is unstable. From observing the graph, it is apparent that there is evidence of structural change. In particular, there are signs of alterations to the values of the parameters as the UK economy extricates itself from the financial crisis and the ensuing period of economic austerity, and steady growth is resumed in the consumption of durable goods.

The lack of temporal stability that has been identified in the relationship between the growth of the consumption of durable goods and the GfK index would seem to suppress any optimism when embarking upon a post-sample analysis to evaluate the predictive performance of the measure of consumer sentiment. Nevertheless, this paper proceeds by comparing the accuracy of out-of-sample forecasts that are generated by models including and excluding the confidence variable. One-step-ahead predictions are produced recursively over the interval, 2006Q3–2016Q3, on the basis that this amounts to approximately one-third of the length of the full sample period. The empirical results that were presented earlier in this section provide a very clear indication that the augmented version of Equation (2) represents an overspecification and so is unlikely to yield satisfactory forecasts. Hence, there is a preference for operating within a much more restricted framework. To be more specific, consideration will be given to whether a regression equation for $\Delta\log(DUR_t)$ which incorporates GfK_{t-1} and GfK_{t-2} (in addition to a constant term) is capable of achieving greater accuracy than a function that omits these terms.[10,11] The somewhat minimalist context within which the predictive credentials of the GfK index are being examined

[10]The choice of two quarterly lags on the GfK index is governed by a comparison of values of the Akaike Information Criterion, having employed as an estimation period, 1986Q2–2006Q2.

[11]Of course, in the absence of the past values of the sentiment measure, log(DUR) is behaving in accordance with a random walk with drift process.

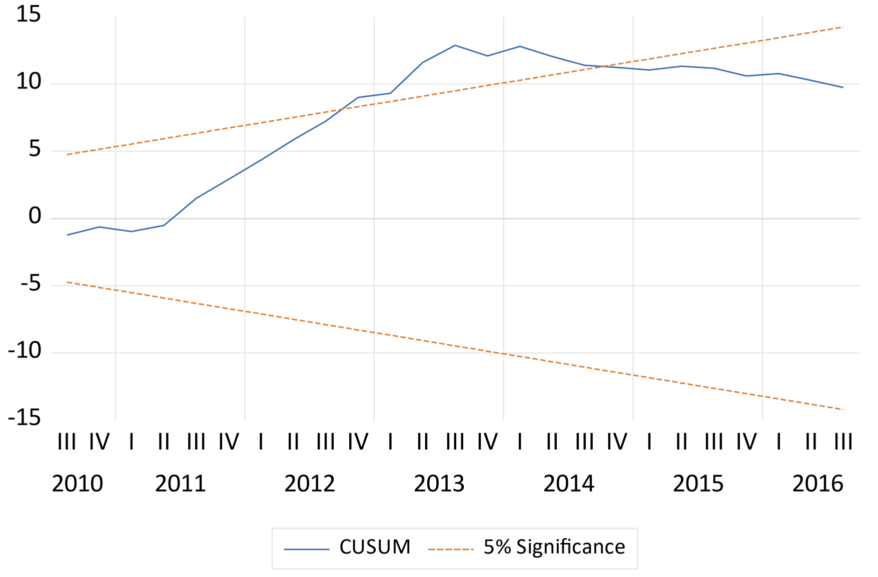

FIGURE 1 Computed values of the CUSUM test statistic corresponding to the augmented version of Equation (2)

can be defended on two grounds. First, results which are not reported but are available from the corresponding author reveal that the full version of Equation (2) and a more parsimonious form of this model generate far less accurate forecasts than the equation involving only GfK_{t-1} and GfK_{t-2}. Second, it seems reasonable to assume that the principal concern of the general public and media is with the ability of consumer confidence 'individually' to deliver advanced warning of developments in household spending.

A familiar approach towards assessing predictive performance is to compare values of a root mean square prediction error (RMSPE) statistic. However, there is also the desire to conduct a formal test of equal forecast accuracy. In spite of reservations about its statistical properties, Bram and Ludvigson (1998) elected to apply a modified version of the Diebold and Mariano (1995) test. On the basis that there are two rival specifications, models i and j, the latter involves the computation of the statistic:

$$S_1 = \frac{\bar{d}_{ij}}{\hat{\sigma}_{\bar{d}_{ij}}}$$

where $\bar{d}_{ij}$ denotes the average loss differential and $\hat{\sigma}_{\bar{d}_{ij}}$ signifies the associated standard error, which requires a robust calculation.[13] Given the manner of its construction, S_1 appears to have asymptotically a standard normal distribution, yet it needs to be respected that, for models with estimated parameters, such a density function only applies if the equations are non-nested (Clark & McCracken, 2001). However, the simulation results of Clark and McCracken (2015) suggested that, when one of the equations is nested within the other, the standard normal critical values

[13]In the forthcoming analysis, $\bar{d}_{ij}$ is calculated from the average, over the forecast interval, of the difference between the squared prediction errors corresponding to the estimated versions of models i and j. A positive value is indicative of the unrestricted equation generally producing more accurate forecasts.

could still be of relevance were there to be a preparedness to depart from the sharp null hypothesis that the expected loss differential is equal to zero. Within the more general specification, when allowance was made for the influence of the additional variables to be non-zero, yet sufficiently small for, in a finite sample, the restricted and unrestricted equations to be expected to generate equally accurate forecasts, the size of the Diebold-Mariano test was typically found to be near to the true level of significance. This result encouraged Diebold (2015) to conclude that, given a desire to perform pseudo out-of-sample model comparisons, traditional Diebold-Mariano tests, utilising $N(0, 1)$ critical values, are likely to be fine.

The modification which Bram and Ludvigson (1998) sought to apply to the Diebold-Mariano test followed the recommendation of Harvey et al. (1997). The latter had observed the Diebold-Mariano test to be quite seriously over-sized in small and moderate samples. In an attempt to remedy this problem, they advocated transforming the statistic, S_1, and basing the test upon a t distribution, rather than the standard normal distribution. To be more specific, in the case one-step-ahead forecasts, Harvey et al. (1997) preferred to calculate the value of:

$$S_1^* = \sqrt{\frac{p-1}{p}}\, S_1$$

where p denotes the number of predictions. Subsequently, the computed value of the statistic should be compared with a critical value corresponding to the t_{p-1} distribution.

Table 3 shows the values of the RMSPE for the two rival equations which have been selected to forecast the values of $\Delta\log(DUR_t)$ over the interval, 2006Q3–2016Q3. In order to provide these figures with some perspective, the average absolute proportional change in spending on durable goods over the 41 quarters is equal to 0.0222. It is apparent that the two values of the RMSPE are almost identical, suggesting the most modest of gains in predictive accuracy as a consequence of utilising data on consumer confidence. The value of the statistic, S_1^*, is only 0.0344, which is associated with a probability value of 0.9727, assuming that a two-tailed test of equal forecast accuracy is being performed.

4 | A DISAGGREGATED ANALYSIS OF UK HOUSEHOLD CONSUMPTION EXPENDITURE

In the introduction to this paper, three criteria were presented which needed to be satisfied in order for a measure of consumer confidence to qualify as a reliable predictor of household consumption expenditure. From the results which were reported in Section 3, pertaining to spending on durable goods, it is apparent that only one out of the three conditions was met. It may be recalled that evidence was obtained of temporal instability in the relationship between the growth of durable goods' consumption and the GfK index. Also, following an out-of-sample analysis, it was discovered that the involvement of past values of the sentiment indicator did not succeed in producing significantly more accurate forecasts than the use of a time-evolving sample mean.[12] However, these findings may not necessitate a refutation of Katona's adaptive theory of consumption. It should be respected that Katona (1968) argued in favour of an attitudinal variable entering a model of 'discretionary' household expenditure. Hence, if there are elements of durable goods'

[12] When a regression equation includes merely a constant term on its right-hand side, the Ordinary Least Squares estimator of the intercept parameter is the same as the sample mean of the dependent variable.

TABLE 3 Values of the RMSPE derived from recursive forecasts of $\Delta\log(DUR_t)$, t = 2006Q3, 2006Q4, ………, 2016Q3

Restricted equation	Unrestricted equation
0.0263	0.0262

Note: The unrestricted equation features on its right-hand side GfK_{t-1} and GfK_{t-2} (in addition to a constant term). When appropriate, the equation also includes the two dummy variables to represent the effect of the Labour Government's vehicle scrappage scheme. The restricted equation is exactly the same, but for the omission of GfK_{t-1} and GfK_{t-2}.

TABLE 4 Examples of discretionary and non-discretionary goods and services

	Discretionary	Percentage change	Non-discretionary	Percentage change
Durable goods	Major durables for outdoor recreation	789.7 (309.5)	Carpets and other floor coverings	40.08 (−35.53)
Semi-durable goods	Glassware, tableware and household utensils	241.1 (57.02)	Books	54.75 (−28.77)
Non-durable goods	Gardens, plants and flowers	201.7 (38.88)	Milk, cheese and eggs	17.22 (−46.04)
Services	Insurance connected with transport	215.5 (45.21)	Hairdressing salons and personal grooming establishments	15.85 (−46.67)

Note: The table shows the percentage change in household consumption expenditure on the respective good or service over the period, 1985Q1–2016Q3. The bracketed figures show the corresponding percentage changes in consumption, expressed as a ratio of real household disposable income. The data on the variables are seasonally adjusted and in the form of constant prices.

consumption which do not warrant the description of inessential then this may well account for the conclusion having been reached that the GfK index is an unreliable predictor.

In the Office for National Statistics publication, *Consumer Trends*, UK household consumption expenditure is divided between spending on durable goods, semi-durable goods, non-durable goods, and services. Within the table that is specifically devoted to durable goods (DGKS), it is possible to identify 15 different categories of expenditure.[13] For each particular component of durable goods' consumption, sufficient data are available to be able to calculate the percentage change from 1985Q1 to 2016Q3. Within Table 4, there can be seen the resultant values for Major Durables for Outdoor Recreation and Carpets and Other Floor Coverings. It is apparent that the two figures are markedly different, indicating that the increase in expenditure on the former has by far exceeded that on the latter. Indeed, the implication is that spending on Carpets and Other Floor Coverings has been remarkably stable, with a growth rate that does not amount to even half of the corresponding rise in real household disposable income.[14] On this basis, it would seem to be inappropriate to be interpreting expenditure on Carpets and Other Floor Coverings as being discretionary, with the consequence that its inclusion in the data on the aggregate consumption of durable goods has possibly weakened the estimated relationship between $\Delta\log(DUR_t)$ and past values of the GfK index.

[13]The label which is given to this table signifies that the associated quarterly data are seasonally adjusted and contained in the form of constant prices.

[14]This statement is made, assuming no substantial falls in the consumption of this good.

While there are aspects of spending on durable goods that are obligatory, it should be appreciated that there are also elements of expenditure on semi-durable goods, non-durable goods, and services which happen to be gratuitous. Indeed, with reference to the latter three categories of consumption, Table 4 contains examples of items for which expenditure has increased much more rapidly than household income over the interval, 1985Q1–2016Q3. It would seem to follow, then, that if there is an aim of compiling a series on discretionary consumer spending, which would enable a fairer test of Katona's theory, consideration needs to be given to the components of all four of the broad classes of consumption. The relevant tables within *Consumer Trends* are those with the labels, DURKS, SDKS, NDKS and SERKS. Within these tables, there are to be found 111 different forms of household expenditure. For each one of these, the following strategy is employed in order to establish whether or not the consumption is discretionary. The preference is for adopting a regression-based approach to ensure that all of the sample data are making a contribution towards the verdict. To be more specific, the two simple equations, (3) and (4), are estimated over the full sample period.

$$Consumption_t = A_0 + A_1 RHDI_t + \varepsilon_t \tag{3}$$

$$\log\left(Consumption_t\right) = C_0 + C_1 \log\left(RHDI_t\right) + \varepsilon_t \tag{4}$$

$$t = 1985Q1,\ 1985Q2,\ \cdots\cdots\cdots,\ 2016Q3$$

It should be emphasised that neither of these equations is intended to be a fully-fledged model of consumption, and so should not be evaluated in that context. With reference to Equation (3), the intercept parameter has the strict interpretation of the average value of household expenditure when the value of real household disposable income is equal to zero. The belief is held that, for an essential good or service, $A_0 > 0$, as consumers would typically find a way of financing the purchase, even were the value of income to be minimal. With reference to Equation (4), on account of both of the left- and right-hand-side variables being contained in a logarithmic form, the parameter, C_1, has the interpretation of the income elasticity of consumption. In the case of an essential good or service, the value of C_1 is considered to be < 1 for the reason that, during an economic downswing (upswing), expenditure is not expected to fall (rise) to the same extent as income.

Consequently, the decision is taken that should both the estimate of A_0 be less than zero and the estimate of C_1 be greater than 1 then the item of consumption expenditure can be classed as discretionary. With regard to the DURKS table within *Consumer Trends*, it transpires that 9 out of the 15 different types of spending on durable goods satisfy the two criteria.[15] Thus, in the year 2015, out of a total of £112,667 million of purchases of durable goods by households, £57,728 million are interpreted as discretionary.[16] Concerning the consumption of semi-durable goods, 13 out of 15 different components are identified as being discretionary, which corresponds to £113,355 million out of a total expenditure of £121,128 million. For non-durable goods, the fraction is 9 out of 33, resulting in £63,036 million out of £255,347 million; and, for services, the ratio is 14 out of 48, giving rise to £152,358 million out of £684,643 million. In sum, then, 45 out of 111 elements of consumption are deemed to be discretionary, which equates with just under 33 per

[15]Specific information can be obtained from the corresponding author on the allocation of the components of consumption not only for durable goods but also for semi-durable goods, non-durable goods, and services.

[16]This equates with just over 51 per cent. This figure would have been far higher but for expenditure on motor vehicles having an estimated income elasticity of 0.92.

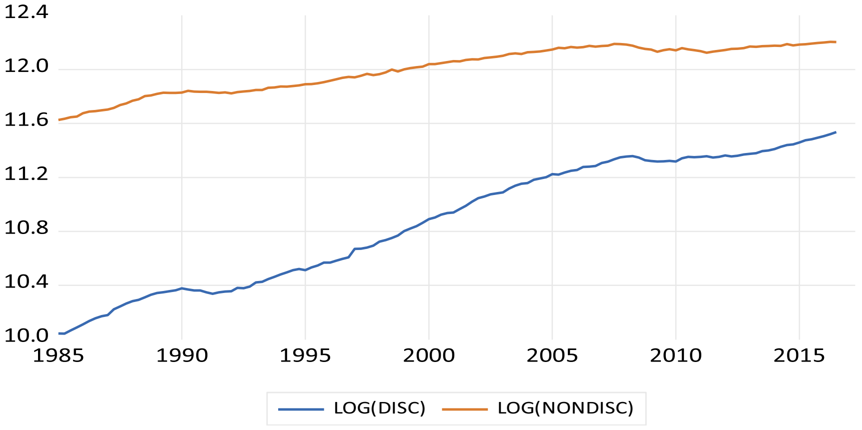

FIGURE 2 Time plots of the logarithms of discretionary and non-discretionary household consumption expenditure in the UK

cent of UK household expenditure in 2015. Interestingly, the contribution that is made by durable goods to overall discretionary spending is only in the region of 15 per cent.

By combining the data on the relevant 45 categories of consumption, it is possible to form a series on the discretionary expenditure by the personal sector (DISC). Also, by subtracting DISC from overall household consumption expenditure, it is possible to achieve a series on aggregate non-discretionary spending (NONDISC). Figure 2 shows the line graphs of the natural logarithms of DISC and NONDISC. It is apparent from a study of this diagram that, over the course of 1985Q1–2016Q3, DISC has fluctuated to a greater extent than NONDISC. More specifically, it can be seen that, in the early 1990s, the fall in NONDISC was much shallower than the decline in DISC. Subsequently, from 1991/1992 to the beginning of the financial crisis. DISC grew noticeably more rapidly than NONDISC. Finally, after the economic recession, from 2011/2012 to the end of the data period, a far stronger recovery can be witnessed in DISC than NONDISC. In a multi-country econometric analysis of consumption behaviour, Gausden and Hasan (2020) remarked on how the ability of a measure of consumer confidence to forecast the growth of household expenditure seemed to be affected by the degree of variability in the latter. Consequently, this may give cause for a degree of optimism when proceeding to investigate the suitability of the GfK index for predicting changes in DISC.

5 | DISCRETIONARY HOUSEHOLD EXPENDITURE AND CONSUMER CONFIDENCE

In the penultimate section of this paper, the empirical methodology which was implemented earlier in connection with the consumption of durable goods will now be applied to discretionary household expenditure. As a consequence of performing this econometric analysis, it will be possible to assess whether the chosen measure of consumer confidence is more suitable for explaining the behaviour of the latter than the former. Indeed, the sentiment variable will be accepted as a reliable leading indicator of discretionary spending if the following three requirements are fulfilled:

TABLE 5 Results obtained from estimating Equations (1) and (2) over the full sample period with $\Delta\log(DISC_{t-i})$ replacing $\Delta\log\left(DUR_{t-i}\right)$, $i = 0, 1, 2, 3, 4$

Right-hand-side variables	Equation (1)	Equation (2)
$\Delta\log\left(DISC_{t-i}\right)$	0.5914 (0.0005)	0.0260 (0.8974)
$\Delta\log\left(RHDI_{t-i}\right)$	0.1147 (0.3460)	−0.0137 (0.4120)
$\Delta\log\left(FTSE_{t-i}\right)$	0.0165 (0.2783)	0.0231 (0.1934)
ΔTB_{t-i}	−0.0046 (0.0116)	−0.0055 (0.2305)
GfK_{t-i}		0.0005 (0.0005)
Adjusted R-squared	0.1719	0.2661
BG(4)	6.7587 (0.1492)	3.2253 (0.5209)

Note: See Table 1.

(i) from having estimated an ARDL model over the full data period, the collective effect of lagged values of the GfK index on the quarterly growth of DISC is statistically significant;
(ii) this statistical significance is preserved when estimating the same equation over sub-periods, and there is no indication of temporal instability;
(iii) from having conducted a post-sample exercise, the involvement of past values of the GfK index leads to a significant improvement in the accuracy of forecasts of $\Delta\log(DISC)$.

Importantly, respecting the overriding objective of this paper, should these three conditions be satisfied then the empirical evidence can be interpreted as providing support for Katona's adaptive theory.

Implementation of the methodology necessitates initially estimation of Equations (1) and (2) over the full data period, but with $\Delta\log(DISC_{t-i})$ replacing $\Delta\log(DUR_{t-i})$ $i = 0, 1, 2, \cdots\cdots\cdots, n$. Having set the number of lags on the variables equal to 4, the results of the estimation and subsequent statistical tests are shown in Table 5. A study of this table reveals that the addition to Equation (1) of the four lags on the GfK index succeeds in raising the value of the adjusted R-squared statistic by 0.0942. Also, within Equation (2), the overall estimated effect of the past values of the confidence measure is positive (0.0005), and so accords with expectations. Furthermore, the GfK index is the only one of the right-hand-side variables for which the F probability value is below a conventional level of significance.'[17][18]

Within the Appendix to this paper, there can also be found Table A5, which shows the consequences of having estimated the augmented version of Equation (2), featuring $\Delta\log\left(NONDISC_{t-i}\right)$ rather than $\Delta\log\left(DUR_{t-i}\right)$, $i = 0, 1, 2, 3, 4$. What we can glean from this table is that the collective effect of past changes in the GfK index on the growth of non-discretionary household expenditure

[17]A comparison of probability values within Table 5 and Table 1 indicates that the GfK index exerts a stronger influence upon $\Delta\log(DISC_t)$ than $\Delta\log(DUR_t)$.

[18]Table A4 in the Appendix to this paper shows individual parameter estimates.

TABLE 6 Results obtained from estimating Equation (2) over selected sub-periods with $\Delta\log(DISC_{t-i})$ replacing $\Delta\log\left(DUR_{t-i}\right)$, $i = 0, 1, 2, 3, 4$

Sub-period	*F* statistic	Probability value
1986Q2–1993Q4	$F(4, 10) = 2.6651$	0.0951
1994Q1–2001Q2	$F(4, 9) = 3.0203$	0.0777
2001Q3–2008Q4	$F(4, 9) = 2.9433$	0.0823
2009Q1–2016Q3	$F(4, 10) = 10.002$	0.0016

Note: See Table 2. It should be recognised, though, that there is no need for the two dummy variables to be additionally included in the equation, when estimating over the interval, 2009Q1–2016Q3.

is positive. However, at a conventional level of significance, it is not possible to reject the null hypothesis that each of the four parameters which are attached to the lags on the confidence variable is equal to zero. We can also see from this table that past values of the dependent variable exert a much stronger influence than in the corresponding equation for $\Delta\log(DISC_t)$. Indeed, the value of the *F* statistic in the first column exceeds the corresponding five per cent level of significance critical value. Such results are to be expected if we have managed to distinguish successfully between goods and services which are optional and those which are essential.

Having estimated Equation (2) over the full sample period, with $\Delta\log\left(DISC_{t-i}\right)$ replacing $\Delta\log\left(DUR_{t-i}\right),\ i = 0, 1, 2, 3, 4$, we now investigate whether the significant relationship between the growth of discretionary expenditure and the GfK index that was obtained is preserved when estimation is performed over sub-periods. For each of the intervals, 1986Q2–1993Q4, 1994Q1–2001Q2, 2001Q3–2008Q4 and 2009Q1–2016Q3, the computed value of an *F* statistic is produced for the purpose of testing the null hypothesis that $B_{5i} = 0,\ i = 1, 2, 3, 4$. Table 6 shows not only these computed values but also the corresponding marginal levels of significance.

With regard to the results in Table 6, there appears to be some consistency in terms of the significance of the relationship between $\Delta\log(DISC_t)$ and $GfK_{t-i},\ i = 1, 2, 3, 4$. Every one of the four probability values in the final column can be observed to be less than 0.10. In contrast to what was seen in Table 2, the lowest (highest) probability value is associated with the most recent (earliest) sub-period.

As a further check on the stability of Equation (2), with discretionary consumption replacing spending on durable goods, we perform the CUSUM test, which was outlined in section 3. Figure 3 shows, in the form of a line graph, the computed values of the test statistic, together with the five per cent level of significance bands. The inference that is drawn from the diagram is that the parameters of the model are constant over time. Over the first half of the data period, there is little deviation from the expected value of zero. It is only when the financial crisis is encountered that some clear departures are visible, although technically none of these are statistically significant.

Finally, attention turns to the usefulness of the GfK index for the purpose of producing out-of-sample, one-quarter-ahead recursive forecasts of $\Delta\log\left(DISC_t\right)$, $t = 2006Q3, 2006Q4, \cdots\cdots\cdots, 2016Q3$. As before, the adopted approach is to compare the predictive accuracy that is achieved by two rival models: an equation which relates the quarterly growth of consumption to (a constant term and) the previous two values of the sentiment measure; and a restricted version which excludes GfK_{t-1} and GfK_{t-2}. Table 7 incorporates the values of the RMSPE statistic for the two models. Through suitably combining these figures, it can be calculated that the inclusion of the lags on consumer confidence succeeds in reducing the value of the RMSPE by more than 25 per cent. Additionally, from

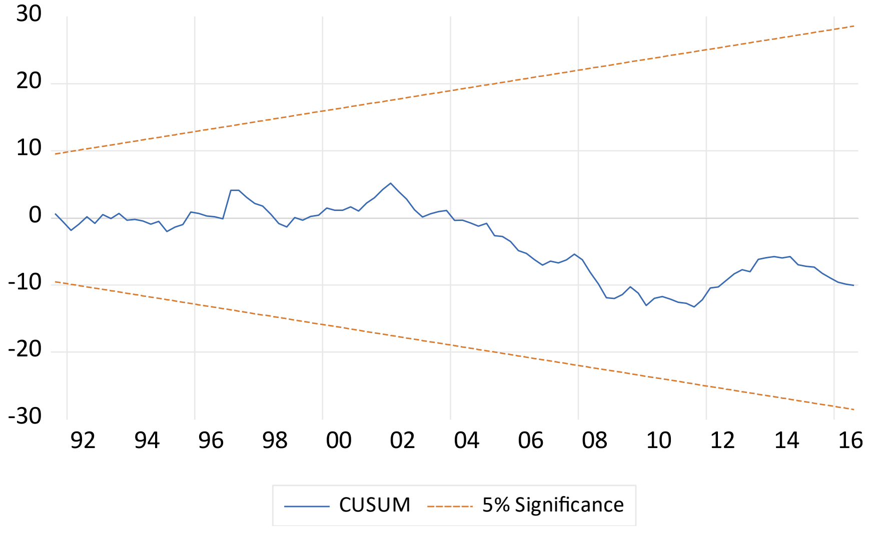

FIGURE 3 Computed values of the CUSUM test statistic corresponding to Equation (2), with $\Delta\log(DISC_{t-i})$ replacing $\Delta\log\left(DUR_{t-i}\right)$, $i = 0,\ 1,\ 2,\ 3,\ 4$

TABLE 7 Values of the RMSPE derived from recursive forecasts of $\Delta\log(DISC_t)$, t = 2006Q3, 2006Q4, ………, 2016Q3

Restricted equation	Unrestricted equation
0.0113	0.0084

Note: The unrestricted equation features on its right-hand side GfK_{t-1} and GfK_{t-2} (in addition to a constant term). The restricted equation is exactly the same, but for the omission of GfK_{t-1} and GfK_{t-2}.

the application of the modified version of the Diebold-Mariano test, which was favoured by Harvey et al. (1997), a value of $S_1^* = 2.0433$ is obtained. For a two-tailed test, founded upon the t_{40} distribution, the corresponding probability value is 0.0476. The latter thereby enables the conclusion to be reached that the involvement in the analysis of past values of the GfK index enables significantly superior forecasts to be produced.

Granted that all three of the stipulated criteria are satisfied, there is justification for maintaining that the opinions and expectations which are formed by households about their own financial position and the general economic situation play a part in deciding the amount of discretionary consumption expenditure that is to be undertaken. Hence, Katona's adaptive theory of consumption would appear to have some credence.

6 | SUMMARY AND CONCLUSION

This paper has had the objective of reappraising Katona’s (1968) adaptive theory of consumer behaviour, which permits a psychological factor a role in determining discretionary household expenditure. In spite of accepting that this may be an imperfect substitute, Katona (1967, 1968) employed spending on durable goods as a proxy for gratuitous consumption. Several more recent econometric studies have shown the relationship between purchases of durable goods and a measure of consumer confidence to be temporally unstable. Such a finding is confirmed in this article, having performed an analysis of UK quarterly data with the GfK index serving as the sentiment indicator.

Rather than regarding these results as evidence which contradicts Katona's theory, this paper questions the suitability of relying upon durable goods expenditure for performing such an investigation and prefers to construct a more genuine representation of discretionary spending. Through undertaking a disaggregated analysis involving the classification of 111 categories of consumption, a time series on non-obligatory expenditure by the household sector is produced (DISC). Following estimation of an established ARDL model over the full data period, 1986Q2–2016Q3, evidence is found of a strongly significant relationship between the growth of DISC and past values of the GfK index. Significance is maintained when the same equation is estimated over four approximately equal length sub-intervals. Also, it is demonstrated that the accuracy of out-of-sample, one-quarter-ahead forecasts of $\Delta\log\left(DISC_t\right)$, $t = 2006Q3,\ 2006Q4,\ \cdots\cdots\cdots,\ 2016Q3$, benefits appreciably from taking into consideration lagged values of the consumer sentiment indicator. Consequently, it appears that, when an evaluation is performed in conjunction with what is believed to be a more relevant consumption variable, the data do not refute Katona's theory.

What we have learnt from our study is that focusing upon the expenditure on durable goods has the effect of downplaying the importance of consumer confidence for stimulating changes in aggregate household consumption. For one reason, it is generally concluded that shifts in attitudes and expectations are of relevance in only limited circumstances. Second, as a proportion of the overall spending by the personal sector in the UK, expenditure on durable goods is far smaller than discretionary consumption. Consequently, we hold the view that decisions on macroeconomic policy would be aided by a reference to the relationship between consumer sentiment and DISC. It should be recognised, though, that regular revisions should be made to the composition of the latter, in accordance with adjustments to tastes and technology. Also, it may be productive to complement the approach that was employed to classify goods and services in this paper with a study of suitable microeconomic data, which show the composition of the consumption of different income groups.

ACKNOWLEDGEMENTS

The authors would like to express their thanks to the Editor of the journal and two anonymous referees for supplying comments which served to enhance the quality of this paper.

REFERENCES

Ahmed, M. I., & Cassou, S. P. (2016). Does consumer confidence affect durable goods spending during good and bad times equally? *Journal of Macroeconomics*, *50*, 86–97. https://doi.org/10.1016/j.jmacro.2016.08.008

Baghestani, H., & Kherfi, S. (2015). An error-correction modeling of U.S. consumer spending: Are there asymmetries? *Journal of Economic Studies*, *42*, 1078–1094. https://doi.org/10.1108/JES-04-2014-0065

Bram, J., & Ludvigson, S. J. (1998). Does consumer confidence forecast household expenditure? A sentiment index horse race. *Federal Reserve Bank of New York Economic Policy Review*, *4*, 59–78.

Brown, R. L., Durbin, J., & Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society: Series B*, *37*, 149–163. https://doi.org/10.1111/j.2517-6161.1975.tb01532.x

Carroll, C. D., Fuhrer, J. C., & Wilcox, D. W. (1994). Does consumer sentiment forecast household spending? If so, why? *The American Economic Review*, *84*, 1397–1408.

Clark, T. E., & McCracken, M. W. (2001). Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics*, *105*, 85–110. https://doi.org/10.1016/S0304-4076(01)00071-9

Clark, T. E., & McCracken, M. W. (2015). Nested forecast model comparisons: A new approach to testing equal accuracy. *Journal of Econometrics*, *186*, 160–177. https://doi.org/10.1016/j.jeconom.2014.06.016

Diebold, F. X. (2015). Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold-Mariano tests. *Journal of Business and Economic Statistics*, *33*, 1–9. https://doi.org/10.1080/07350015.2014.983236

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, *13*, 253–263.

Easaw, J., Garratt, D., & Heravi, S. (2005). Does consumer sentiment accurately forecast UK household consumption? Are there any comparisons to be made with the US? *Journal of Macroeconomics*, *27*, 517–532. https://doi.org/10.1016/j.jmacro.2004.03.001

Easaw, J. Z., & Heravi, S. M. (2004). Evaluating consumer sentiments as predictors of UK household consumption behavior. Are they accurate and useful? *International Journal of Forecasting*, *20*, 671–681. https://doi.org/10.1016/j.ijforecast.2003.12.006

Gausden, R., & Hasan, M. S. (2016). Would information on consumer confidence have helped to predict UK household expenditure during the recent economic crisis? *Applied Economics*, *48*, 1695–1709. https://doi.org/10.1080/00036846.2015.1105926

Gausden, R., & Hasan, M. S. (2020). Comparative performances of measures of consumer and economic sentiment in forecasting consumption: A multi-country analysis. *Applied Economics*, *52*, 1088–1104. https://doi.org/10.1080/00036846.2019.1659489

Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, *13*, 281–291. https://doi.org/10.1016/S0169-2070(96)00719-4

Katona, G. (1967). Anticipations statistics and consumer behavior. *The American Statistician*, *21*, 12–13.

Katona, G. (1968). Consumer behavior: Theory and findings on expectations and aspirations. *The American Economic Review*, *58*, 19–30.

Throop, A. W. (1992). Consumer sentiment: Its causes and effects. *Federal Reserve Bank of San Francisco Economic Review*, *1*, 35–59.

How to cite this article: Gausden, R., & Hasan, M. (2022). A reappraisal of Katona's adaptive theory of consumer behaviour using U.K. data. *The Manchester School*, 00, 1–22. https://doi.org/10.1111/manc.12395

APPENDIX

1 | DATA SOURCES

Household Consumption Expenditure
Source: Consumer Trends database of the Office for National Statistics, Quarter 2, 2017
Quarterly, seasonally-adjusted data in the form of constant (2015) prices
Real Household Disposable Income
Source: Datastream. Codename: NRJR
Quarterly, seasonally-adjusted data
FTSE All Share Price Index
Source: Datastream. Codename: FTALLSH(PI)
Data are quarterly averages, seasonally unadjusted
Implicit Price Deflator

The implicit price deflator is achieved by dividing current-price consumer spending (domestic) by its constant-price counterpart Source: Datastream. Codenames: ZAKV; ZAKW
Quarterly, seasonally-adjusted data
FTSE All Share Price Index (real terms)
The series on the FTSE All Share Price Index, in real terms, is created by dividing the nominal variable by the implicit price deflator
Rate of Discount, Three-Month Treasury Bill
Source: Bank of England database. Codename: IUQAAJNB
Data are quarterly averages, seasonally unadjusted
The GfK Consumer Confidence Indicator
Source: European Commission Consumer Survey Data https://ec.europa.eu/info/business-economy-euro/indicators-statistics/economic-databases/business-and-consumer-surveys/download-business-and-consumer-survey-data/time-series_en Data are quarterly averages, seasonally adjusted

TABLE A1 Results obtained from estimating augmented Equation (2) over the full sample period

	Right-hand-side variable				
Lag length	$\Delta\log(DUR_{t-i})$	$\Delta\log(RHDI_{t-i})$	$\Delta\log(FTSE_{t-i})$	ΔTB_{t-i}	GfK_{t-i}
$i = 1$	−0.2577***	0.2491	0.0394	−0.0046	0.0004
$i = 2$	0.0527	−0.0020	−0.0036	0.0029	0.0007
$i = 3$	0.0516	−0.4751*	0.0383	−0.0013	0.0008
$i = 4$	0.1457	−0.1282	0.0628**	0.0023	−0.0009
Sum	−0.0077	−0.3562	0.1370	−0.0008	0.0010
$F(4, 99)$ (prob.)	3.4124** (0.0117)	1.2269 (0.3043)	1.8881 (0.1185)	0.7145 (0.5839)	3.4875** (0.0104)

Note: The equation has been estimated by Ordinary Least Squares over the sample period, 1986Q2–2016Q3. Tests of hypotheses are performed with allowance for heteroskedasticity and autocorrelation in the disturbance terms (via selecting the Newey-West covariance procedure). The equation incorporates additionally two dummy variables to represent the redistribution of expenditure which occurred over 2009 and 2010 as a consequence of the implementation of the Labour Government's vehicle scrappage scheme. The figures in the first four rows are the estimates of the parameters which are attached to the lags on the variables. The figures in the fifth row are the sums of the respective four estimates. In the final row, each of the computed values of the F statistic corresponds to a test of the joint null hypothesis that all four parameters are equal to zero. The figures in brackets are the associated probability values. Significance at the 1, 5 and 10 per cent levels is denoted by ***, ** and *, respectively.

TABLE A2 Results obtained from estimating augmented Equation (2) over the full sample period with CCI replacing GfK as the confidence indicator

	Right-hand-side variable				
Lag length	$\Delta\log(DUR_{t-i})$	$\Delta\log(RHDI_{t-i})$	$\Delta\log(FTSE_{t-i})$	ΔTB_{t-i}	CCI_{t-i}
$i = 1$	−0.2203**	0.3156	0.0386	−0.0049	0.0008
$i = 2$	0.0842	0.1223	−0.0003	0.0020	−0.0002
$i = 3$	0.0820	−0.3768	0.0401	−0.0031	0.0007
$i = 4$	0.1756*	−0.0433	0.0682**	0.0007	−0.0005
Sum	0.1216	0.0178	0.1465	−0.0053	0.0008
$F(4, 99)$ (prob.)	3.6540*** (0.0081)	1.0095 (0.4063)	2.0428* (0.0942)	0.6537 (0.6256)	1.7558 (0.1439)

Note: See Table A1.

TABLE A3 Results obtained from estimating augmented Equation (2) over the full sample period with MORI replacing GfK as the confidence indicator

	Right-hand-side variable				
Lag length	$\Delta\log(DUR_{t-i})$	$\Delta\log(RHDI_{t-i})$	$\Delta\log(FTSE_{t-i})$	ΔTB_{t-i}	$MORI_{t-i}$
$i = 1$	−0.1921*	0.3805	0.0546	−0.0019	0.0003
$i = 2$	0.1311*	0.1940	−0.0010	0.0024	−0.0006
$i = 3$	0.1024	−0.4067	0.0420	−0.0026	0.0006
$i = 4$	0.2563**	−0.1840	0.0669**	0.0002	0.0001
Sum	0.2978	−0.0162	0.1626	−0.0019	0.0004
$F(4, 99)$ (prob.)	5.1851*** (0.0008)	1.1362 (0.3440)	2.1878* (0.0758)	0.1091 (0.9791)	1.6596 (0.1654)

Note: See Table A1.

TABLE A4 Results obtained from estimating Equation (2) over the full sample period with $\Delta\log(DISC_{t-i})$ replacing $\Delta\log\left(DUR_{t-i}\right)$, $i = 0, 1, 2, 3, 4$

	Right-hand-side variable				
Lag length	$\Delta\log(DISC_{t-i})$	$\Delta\log(RHDI_{t-i})$	$\Delta\log(FTSE_{t-i})$	ΔTB_{t-i}	GfK_{t-i}
$i = 1$	−0.0177	−0.1474	0.0159	−0.0022	0.0002
$i = 2$	0.0269	0.0126	0.0139	0.0004	0.0002
$i = 3$	0.0810	0.0688	−0.0069	−0.0023	0.0004
$i = 4$	−0.0643	0.0523	0.0002	−0.0015	−0.0003
Sum	0.0260	−0.0137	0.0231	−0.0055	0.0005
$F(4, 101)$ (prob.)	0.2689 (0.8974)	0.9987 (0.4120)	1.5505 (0.1934)	1.4269 (0.2305)	5.4381*** (0.0005)

Note: See Table A1. It should be emphasised, though, that, when $\Delta\log(DISC_t)$ operates as the dependent variable, there is no need for the equation to incorporate the two dummy variables.

TABLE A5 Results obtained from estimating augmented Equation (2) over the full sample period with $\Delta\log(NONDISC_{t-i})$ replacing $\Delta\log\left(DUR_{t-i}\right)$, $i = 0, 1, 2, 3, 4$

	Right-hand-side variable				
Lag length	$\Delta\log(NONDISC_{t-i})$	$\Delta\log(RHDI_{t-i})$	$\Delta\log(FTSE_{t-i})$	ΔTB_{t-i}	GfK_{t-i}
$i = 1$	−0.0877	0.1025*	0.0031	−0.0006	0.0005*
$i = 2$	0.1919**	−0.0548	−0.0084	0.0022**	−0.0003
$i = 3$	0.2381**	−0.0553	0.0096	−0.0010	−0.0001
$i = 4$	0.1241	−0.0296	0.0069	−0.0010	0.0001
Sum	0.4664	−0.0371	0.0113	−0.0004	0.0002
$F(4, 99)$ (prob.)	2.7136** (0.0342)	2.9489** (0.0238)	1.0236 (0.3990)	1.7589 (0.1432)	1.4722 (0.2164)

Note: See Table A1.