

# Consumer Sentiment and Consumer Spending: Decomposing the Granger Causal Relationship in the Time Domain

S. GELPER\*

A. LEMMENS

C. CROUX

K.U.Leuven

## Abstract

It is often believed that the consumer sentiment index has predictive power for future consumption levels. While Granger causality tests have already been used to test for this, no attempt has been made yet to quantify the predictive power of the consumer sentiment index over different time horizons. In this paper, we decompose the Granger causality at different time lags, by looking at a sequence of nested prediction models. Since the consumer sentiment index turns out to be cointegrated with real consumption, we resort to Error Correcting Models. Four consumption series are studied, namely total real consumption, real consumption of durables, nondurables and services. Among other findings, we show that the consumer sentiment index Granger causes future consumption with an average time lag of four to five months. Furthermore, it is found that the consumer sentiment index has more incremental predictive power for consumption of services than for consumption of durables or nondurables, and that the index is not only useful as a predictor at the very short term, but keeps predictive power at larger time horizons.

*Keywords:* Cointegration, Consumer Sentiment, Granger Causality, VEC models

*JEL classification:* C32; E21

---

\*Corresponding author: Sarah Gelper, Dept. Applied Economics, K.U.Leuven, Naamsestraat 69, B-3000 Leuven. e-mail: sarah.gelper@econ.kuleuven.be

# 1 Introduction

In this article, we investigate the extent to which the US Consumer Sentiment Index (CSI) offers relevant and timely insights into future real consumption levels in the United States. The basic idea behind this research question is that if US consumers feel confident about the actual and future economic and financial situation, they would be more willing to increase their consumption. In contrast, pessimistic consumers could theoretically save more money and delay their spending further in time. This problem has already been investigated by several authors and the results were rather divided (see also Ludvigson, 2004 and Vuchelen, 2004). On the one hand, Desroches and Gosselin (2002) and Roberts and Simon (2001) found that sentiment indicators contain little information to forecast consumption, even if Desroches and Gosselin (2002) stressed the importance of the consumer sentiment in times of high economic and political uncertainty. On the other hand, Batchelor and Dua (1998), Carroll, Fuhrer and Wilcox (1994), Easaw and Heravi (2004), Eppright, Arguea and Huth (1998), Huth, Eppright and Taube (1994), Kumar et al. (1995) and Souleles (2004) all found the consumer sentiment index to be a useful leading indicator in predicting aggregate consumer expenditures. In this paper we want to study the term structure of the consumer sentiment index. Nobody, up to our knowledge, has looked at the time lag at which the consumer sentiment helps to explain consumption. Is the predictive power of the index mainly present at the very short run, or is it pertinent for larger forecasting horizons?

The research question to which extent the US consumer sentiment index offers relevant and timely insights into future real consumption levels in the United States, can be translated into a Granger causality framework. One variable is said to Granger cause the other if it helps to make a more accurate forecast of the other variable than had we only used the past of the latter as predictor. Note that Granger causality between two variables can not be interpreted as a real causal relationship but merely shows that one variable is leading the other one. The first research question becomes now “Does the US consumer sentiment index Granger cause US real consumption, and/or any of its components?”. Most previous studies used Vector Autoregression (VAR) models to investigate this question. For example Utaka (2003) models Japanese consumer confidence and GDP as a VAR and finds that the confidence indicator Granger causes GDP, at least when working with monthly or quarterly data. However, US

consumer confidence and consumption turn out to be cointegrated (as was also found by Throop, 1992). Consequently, a Vector Error Correcting model is more appropriate.

In a second part of the analysis, we compute a measure of Granger causality which allows to separate the Granger causal effect at different time lags. Indeed, when making the prediction for next month consumption level, not only the most recent value of the consumer index is relevant, but also past values. The importance of the consumer sentiment index is then measured at different lags. As such, the Granger causal relationship can be decomposed over the time domain.

In the third and last part of this empirical study, it will be shown that the information in present and past values of the CSI is not only of interest for making predictions at very short time horizons, but still has predictive power for larger forecasting horizons. Forecasts are again made with an Error Correcting model, taking the long run relationship between CSI and consumption into account.

Real consumption data are decomposed in three important parts, i.e. durables, non-durables and service consumption. As the consumer's decision on purchasing durables, non-durables and services are led by different motives, one may expect that the consumer sentiment index does not equally affect all consumption components. For instance, durable goods' consumption is known to be more easily delayed in time than non-durables'. Therefore, the analysis is made for each of the consumption components separately. The heterogeneity of the Granger causal effect on different components of consumer sentiment spending was also studied by Throop (1992) and Ludvigson (2004), but they considered a different decomposition of personal consumption.

The remainder of this paper is organized as follows. Section two covers the methodological aspect of testing for Granger causality and its decomposition over the time domain. The decomposition of the Granger causality measure for a vector autoregressive model is defined in Gouriéroux and Monfort (1990). We extend this approach to cointegrated time series using nested Vector Error Correcting models. In the third section, the data are described. Section four contains the empirical results. In particular, it is investigated at what time lag the consumer sentiment index offers valuable information about consumption. A graphical

representation of the results is provided. Furthermore, we examine the forecasting power of the consumer sentiment for consumers' spending at different forecast horizons. Finally, section five concludes.

## 2 Methodology

This section describes the methodology which will be used to investigate the Granger causal relationship between U.S. consumer sentiment and real US consumer spending. Recall that time series  $X_t$  Granger causes a time series  $Y_t$  if the past of  $X_t$  helps to forecast the future of  $Y_t$  after controlling for the past of  $Y_t$ .

The Vector AutoRegression (VAR) framework allows to test for Granger causality and explicitly includes the possibility of a feedback causality. For  $X_t$  and  $Y_t$  two stationary time series, a bivariate VAR model of order  $M$  is given by

$$X_t = \alpha_1 + \sum_{k=1}^M \beta_{1,k} X_{t-k} + \sum_{k=1}^M \gamma_{1,k} Y_{t-k} + \varepsilon_t^x, \quad (2.1)$$

$$Y_t = \alpha_2 + \sum_{k=1}^M \beta_{2,k} Y_{t-k} + \sum_{k=1}^M \gamma_{2,k} X_{t-k} + \varepsilon_t^y, \quad (2.2)$$

where the error terms  $\varepsilon_t^x$  and  $\varepsilon_t^y$  are assumed to be Gaussian white noise with zero mean and a constant covariance matrix. After estimating equations (2.1) and (2.2), several tests for Granger causality can be conducted. The series  $X_t$  Granger causes  $Y_t$  if the  $\gamma_{2,k}$  coefficients are jointly significant, while  $Y_t$  Granger causes  $X_t$  if the  $\gamma_{1,k}$ 's are jointly significant. If both the  $\gamma_{1,k}$  coefficients and the  $\gamma_{2,k}$  coefficients are jointly significant, there is evidence for a feedback relationship between  $X_t$  and  $Y_t$ .

### 2.1 Decomposition in the time domain

As previously mentioned, we are especially interested in the decomposition of the Granger causality of  $X_t$  for  $Y_t$  over different time lags. We follow a procedure proposed by Gouriéroux and Monfort (1990) to determine the contribution of each time lag to the strength of the

Granger causal relationship. Consider the following pair of equations:

$$Y_t = \alpha + \sum_{k=1}^M \beta_k Y_{t-k} + \sum_{k=j+1}^M \gamma_k X_{t-k} + \varepsilon_t^{j+1}, \quad (2.3)$$

and

$$Y_t = \alpha + \sum_{k=1}^M \beta_k Y_{t-k} + \sum_{k=j}^M \gamma_k X_{t-k} + \varepsilon_t^j, \quad (2.4)$$

which both can be estimated by maximum likelihood for  $j = 1, \dots, M$ . Since the only difference between (2.3) and (2.4) is an additional regressor in (2.4), namely  $X_{t-j}$ , model (2.3) is nested in model (2.4). As a consequence, the estimated variance of the error terms in equation (2.4),  $\hat{\sigma}_j^2$ , will always be smaller than or equal to the error term variance for equation (2.3),  $\hat{\sigma}_{j+1}^2$ . If the difference is large, the variable  $X_{t-j}$  has significant power when forecasting  $Y_t$  and controlling for its previous values. The measure of causality at lag  $j$  is then defined as

$$C_j = \ln \frac{\hat{\sigma}_{j+1}^2}{\hat{\sigma}_j^2}. \quad (2.5)$$

The causality measure at lag  $j$  describes the forecasting power of  $X_{t-j}$  for  $Y_t$ , after controlling for the past of  $X_{t-j}$  and the past of  $Y_t$ . As  $\hat{\sigma}_{j+1}^2 \geq \hat{\sigma}_j^2$ , it follows that  $C_j \geq 0$ , resulting in a measure for the strength of  $X_{t-j}$  in Granger causing  $Y_t$ .

Once  $C_j$  is computed, we would like to know whether it is significantly different from zero. The critical values of  $C_j$  can be derived from its resemblance with the likelihood ratio (LR) test statistic. This test can be used to compare the nested models (2.3) and (2.4) and tests for the null hypothesis that  $\gamma_j$  equals zero. If we call (2.4) the unrestricted model and (2.3) the restricted model, then the LR test statistic takes the form

$$LR = 2 \left( \log L(\theta_U) - \log L(\theta_R) \right). \quad (2.6)$$

Here  $\log L(\theta_U)$  denotes the loglikelihood at the unrestricted model, with  $\theta_U$  the parameter vector collecting the estimate of all  $\beta_k$ , all  $\gamma_k$  and of the variance of the error terms. Analogously for the loglikelihood function at the restricted model,  $\log L(\theta_R)$ . It is not difficult to show, e.g. as in Gouriéroux and Jasiak (2001, chapter 4) that

$$LR = T \times C_j \sim \chi_1^2,$$

where  $T$  is the number of observations. Hence, if  $C_j$  is larger than the critical value  $\chi_{1,1-\alpha}^2/T$  (with  $\chi_{1,1-\alpha}^2$  the  $\alpha$ -upper quantile of the chi-squared distribution with one degree of freedom)

then the Granger causality measure at lag  $j$  is found to be asymptotically significant at level  $\alpha$ .

When summing up all  $C_j$ , for  $j$  ranging from 1 to  $M$ , a total measure of causality is obtained as

$$C = \sum_{j=1}^M C_j = \ln \frac{\hat{\sigma}_{M+1}^2}{\hat{\sigma}_1^2}. \quad (2.7)$$

It measures the total effect of all the lagged  $X_t$  on the present  $Y_t$ . Indeed, for  $j = M$ , equation (2.3) reduces to a regression with only lagged variables  $Y_t$  on the right hand side, and we call this the empty model. On the other hand, for  $j = 1$ , we call equation (2.4) the full model since the whole past of  $X_t$  is included. The total measure of causality  $C$  compares the error-variance in the empty and the full model and is again closely related to a likelihood ratio test for  $H_0 : \gamma_1 = \gamma_2 = \dots = \gamma_M = 0$  in the full model. In fact, we have that  $T \times C$  equals the Granger-Wald test for Granger causality. Note that this causality measure  $C$  was initially proposed by Gouriéroux, Monfort and Renault (1987) on the basis of the Kullback Information Criterion.

An interesting summary measure that can be derived from the causality measures at different lags is the *mean causality lag* at which the lagged values of  $X_t$  help to forecast  $Y_t$ . This measure is denoted by  $D$  and computed as the average of the lag orders weighted by the causality measures at each lag:

$$D = \sum_{j=1}^M \frac{j C_j}{C}. \quad (2.8)$$

The mean causality lag  $D$  indicates how long a value of the CSI remains pertinent. For example, if  $D$  is small, then only the most recent values of the CSI are important for predicting future consumption. On the other hand, if  $D$  is large, then the consumers' spending is reflected in a much longer series of past CSI values.

## 2.2 The Vector Error Correcting Model

The procedure proposed by Gouriéroux and Monfort (1990) for decomposing Granger causality over the time domain is attractive, but requires both series to be stationary. If  $X_t$  and  $Y_t$  are non-stationary, a VAR model applied on the series in differences could be taken.

In our setting, however, it turns out that the consumer sentiment index is nonstationary but also *cointegrated* with total real consumption, as well as with each of its three components. As we know, two time series are cointegrated if there is a long run relationship between them. If we want to take this long run effect into account, an error correcting term has to be included in the model. This allows to separate the short run from the long run causality. Therefore, the Vector Error Correcting (VEC) model is used:

$$\Delta X_t = \alpha_1 + \sum_{k=1}^M \beta_{1,k} \Delta X_{t-k} + \sum_{k=1}^M \gamma_{1,k} \Delta Y_{t-k} + \delta_1 U_{t-1} + \varepsilon_t^x \quad (2.9)$$

$$\Delta Y_t = \alpha_2 + \sum_{k=1}^M \beta_{2,k} \Delta Y_{t-k} + \sum_{k=1}^M \gamma_{2,k} \Delta X_{t-k} + \delta_2 U_{t-1} + \varepsilon_t^y, \quad (2.10)$$

where  $X_t$  and  $Y_t$  are integrated of order one and cointegrated time series. Equations (2.9) and (2.10) are the same as equations (2.1) and (2.2), but an *error correction term*, denoted by  $U_{t-1}$ , is added. This lagged error correction term is obtained from the Johansen procedure (Johansen, 1995). This procedure estimates by maximum likelihood the long run relationship as a cointegration equation

$$U_t = aX_t + bY_t + c, \quad (2.11)$$

with  $a = 1$  by normalization. The error correcting term  $U_{t-1}$  gives the deviation from the long term equilibrium at  $t - 1$ , and will lead to a short run adjustment in  $X_t$ ,  $Y_t$  or both. The coefficients  $\delta_1$  and  $\delta_2$  in (2.9) and (2.10) measure the speed of adjustment. If the series  $X_t$  and  $Y_t$  are cointegrated, then  $\delta_1$  and  $\delta_2$  are jointly significant, due to the representation theorem of Engle and Granger (1987). Since we want to measure Granger causality of  $X_t$  for  $Y_t$ , we focus on equation (2.10) in the sequel.

As explained in more detail in Miller and Russek (2001), the null hypothesis that  $X_t$  does not Granger cause  $Y_t$  is rejected not only if the  $\gamma_{2,k}$ 's in equation (2.10) are jointly significant, but also when only  $\delta_2$  is significant. The coefficients  $\gamma_{2,k}$  give insight about the Granger causality between the short run components of the time series. If the  $\gamma_{2,k}$ 's are found to be significant, it implies that the *changes* in  $X_t$  are important in predicting future  $Y_t$ . The coefficient  $\delta_2$  provides evidence about the Granger causality between the deviation from the long run equilibrium and  $Y_t$ . If  $\delta_2$  is significant, the *level* of  $X_t$  is important in predicting future  $Y_t$ .

In the cointegration framework, the method of Gouriéroux and Monfort (1990) for decomposing Granger causality needs to be modified slightly. For measuring the causality at lag  $j$ , we consider now the two models

$$\Delta Y_t = \alpha + \sum_{k=1}^M \beta_k \Delta Y_{t-k} + \sum_{k=j+1}^M \gamma_k \Delta X_{t-k} + \delta U_{t-(j+1)} + \varepsilon_t^{j+1}, \quad (2.12)$$

and

$$\Delta Y_t = \alpha + \sum_{k=1}^M \beta_k \Delta Y_{t-k} + \sum_{k=j}^M \gamma_k \Delta X_{t-k} + \delta U_{t-j} + \varepsilon_t^j. \quad (2.13)$$

Because we want model (2.12) to take only  $X_{t-j-1}$  and past values into account, we include the deviation from the long run equilibrium in period  $t - (j + 1)$ , that is  $U_{t-(j+1)}$ . For the same reason, we include  $U_{t-j}$  in equation (2.13). Equations (2.12) and (2.13) are both estimated by maximum likelihood. For the error correcting term, we use  $X_{t-j} + \hat{b}Y_{t-j} + \hat{c}$  as estimate for  $U_{t-j}$  (and similarly for  $U_{t-(j+1)}$ ), where  $\hat{b}$  and  $\hat{c}$  are obtained by preliminary maximum likelihood estimation of the cointegration equation. The latter estimates are super consistent, i.e. converge at order  $1/T$ . The estimates for the parameters in models (2.12) and (2.13) converge at the normal, slower, rate of  $1/\sqrt{T}$ . Hence, when performing inference for the parameters  $\gamma_k$  and  $\delta$  in (2.12) and (2.13), the parameters  $a$ ,  $b$  and  $c$  hidden in the definition of  $U_{t-j}$  may be considered as constant and given by their preliminary estimates. A formal proof of this argument is given in Toda and Phillips (1994).

Again, we expect the estimated variance of  $\varepsilon_t^j$ ,  $\hat{\sigma}_j^2$ , to be smaller than or equal to the estimated variance of  $\varepsilon_t^{j+1}$ ,  $\hat{\sigma}_{j+1}^2$ . Herefore, it is sufficient to show that model (2.12) is nested in model (2.13). Rewrite model (2.12) as

$$\Delta Y_t = \alpha + \sum_{k=1}^M \beta_k \Delta Y_{t-k} + \sum_{k=j+1}^M \gamma_k \Delta X_{t-k} + \gamma_j \Delta X_{t-j} + \delta U_{t-j} + \varepsilon_t. \quad (2.14)$$

The fourth and fifth term of the right hand side of this equation can be rewritten as

$$\begin{aligned} & \gamma_j \Delta X_{t-j} + \delta U_{t-j} \\ &= \gamma_j \Delta X_{t-j} + \delta(aX_{t-j} + bY_{t-j} + c) - \delta(aX_{t-j-1} + bY_{t-j-1} + c) + \delta(aX_{t-j-1} + bY_{t-j-1} + c) \\ &= \gamma_j \Delta X_{t-j} + \delta a \Delta X_{t-j} + \delta b \Delta Y_{t-j} + \delta(aX_{t-j-1} + bY_{t-j-1} + c) \\ &= (\gamma_j + \delta a) \Delta X_{t-j} + \delta b \Delta Y_{t-j} + \delta U_{t-j-1}. \end{aligned}$$



When plugging this back into the equation (2.14), we get

$$\begin{aligned}\Delta Y_t = & \alpha + \sum_{k=1}^{j-1} \beta_k \Delta Y_{t-k} + (\beta_j + \delta b) \Delta Y_{t-j} + \sum_{k=j+1}^M \beta_k \Delta Y_{t-k} \\ & + \sum_{k=j+1}^M \gamma_k \Delta X_{t-k} + (\gamma_j + \delta a) \Delta X_{t-j} + \delta U_{t-j-1} + \varepsilon_t.\end{aligned}\quad (2.15)$$

This transformation shows that model (2.13) can be rewritten in the form of model (2.12) plus an additional term in  $\Delta X_{t-j}$ . Hence, model (2.12) is nested in model (2.13) and is resulting from it by imposing the restriction that  $\gamma_j + \delta a = 0$ .

The measure of the forecasting power of  $X_{t-j}$  for  $Y_t$ , after controlling for the past of  $X_{t-j}$  and the past of  $Y_t$  is then defined in the same way as in (2.5),

$$C_j = \ln \frac{\hat{\sigma}_{j+1}^2}{\hat{\sigma}_j^2},$$

where the residual variances are now estimated via the VEC approach instead of the VAR model. Because equations (2.12) and (2.13) are nested, the likelihood ratio test is again appropriate for testing whether the Granger causality measure at lag  $j$  is significant. The definitions of the total Granger causality measure  $C$  and the mean Granger causality lag  $D$ , as defined in (2.7) and (2.8), are again applicable.

### 3 Data

The US consumer sentiment index is computed by the Survey Research Center at the University of Michigan<sup>1</sup>. It is derived from the answers to five questions asked to US consumers about (i) the financial situation of households compared to one year ago, (ii) the expected financial situation of households within one year, (iii) the expected general economic/financial situation of the country over the next twelve months, (iv) the economic (unemployment) expectations during the next five years, and (v) the appropriateness of buying major household durables at present. The consumer sentiment index is then computed as a weighted average of the relative scores (percentage of favorable answers minus percentage of negative

---

<sup>1</sup>Data available at <http://www.sca.isr.umich.edu/main.php>

answers, plus 100) for each of the five questions. The consumer sentiment index surveyed at  $t$  will be denoted by  $CSI_t$ . An augmented Dickey-Fuller test for presence of a unit root<sup>2</sup> indicates that the consumer sentiment index series needs to be differentiated of order one ( $\Delta CSI_t = CSI_t - CSI_{t-1}$ ) to become stationary. The consumer sentiment index  $CSI_t$  plays the role of  $X_t$  in the previous section on the methodology, and is aimed at Granger causing real consumption, denoted by  $Y_t$ .

The US real consumption data are collected by the US Bureau of Economic Analysis, an agency of the US Department of Commerce<sup>3</sup>. The data range from January 1978 to February 2004, resulting in 314 observations. These monthly time series are quantity indices, and are seasonally adjusted at annual rates by the data provider. The data measure the goods and services purchased by the persons resident in the United States. Total US real consumption at time  $t$  is denoted by  $RC_t$  and has three major components: (i) consumption of durables ( $RD_t$ ) consisting largely of motor vehicles, furniture and household equipment, (ii) consumption of nondurables ( $RND_t$ ) containing, among others, expenditures on food, clothing and gasoline and (iii) consumption of services ( $RS_t$ ) consisting mainly of rental housing, recreation, medical care and transportation. All four consumption series have been log-transformed and have a stochastic trend<sup>4</sup>.

## 4 Results

Using the Johansen procedure, the consumer sentiment index is found to be cointegrated with total real consumption and each of its three components<sup>5</sup>. Therefore, a Granger-causality

---

<sup>2</sup>The ADF test statistic of order two (intercept in test statistic) equals -2.45 ( $p > 0.10$ ), suggesting no rejection of the null of a unit root.

<sup>3</sup>Data available at <http://www.bea.gov/bea/dn/nipaweb/index.asp>

<sup>4</sup>The ADF test statistic of order two (intercept and trend in test statistic) equals -1.78 ( $p > 0.10$ ) for total consumption, -2.01 ( $p > 0.10$ ) for consumption of durables, -1.05 ( $p > 0.10$ ) for consumption of nondurables and -1.50 ( $p > 0.10$ ) for services consumption, suggesting no rejection of the null of a unit root in all cases.

<sup>5</sup>Using the Johansen trace-test for cointegration (of order two, with trend in the cointegration equation and no deterministic trend in the VAR model), the null of no cointegration between the series of consumer sentiment and the four consumption series is rejected at the 1% level for total real consumption and consumption of durables, and at the 5% level for nondurables and service consumption.

analysis including an error correction term, as described in Section 2.2, is performed to gain insight into the term structure of the underlying causality between the time series. In the error correcting model, equation (2.10) is used to determine whether the consumer sentiment index is Granger causing real consumption or one of its components. Equation (2.9) examines the opposite causality, which is not the topic of interest in this paper. Since the time series  $\log RC_t$ ,  $\log RD_t$ ,  $\log RND_t$  and  $\log RS_t$  all contain a stochastic trend and the  $CSI_t$  does not, a trend variable is included in the cointegration equation (2.11).

#### 4.1 Granger Causality Tests

We estimate the VEC model and test separately for short run, long run and overall Granger causality. The lag length of the VEC model was selected according to the Schwartz criterion. In terms of equation (2.10), short run causality is present if the  $\gamma_{2,k}$ 's are jointly significant, long run causality if  $\delta_2$  is significant, and overall causality if the  $\gamma_{2,k}$ 's and the  $\delta_2$  are jointly significant. Table 1 summarizes the results of the Granger causality tests for each of the four consumption series. The results in Table 1 show that there is strong evidence

Table 1: Test for no Granger causality of the consumer sentiment index for consumption and its components using the VEC model. The test statistic for the long run follows a  $t$ -distribution and for the short run and overall Granger causality an  $F$ -distribution. Corresponding  $p$ -values are reported between parentheses.

	Short run	Long run	Overall
total consumption	1.010 (0.366)	4.207 (<0.001)	11.553 (<0.001)
durables consumption	0.785 (0.457)	3.074 (0.002)	6.331 (0.002)
nondurables consumption	0.565 (0.569)	2.759 (0.006)	4.956 (0.008)
service consumption	2.804 (0.062)	3.771 (<0.001)	12.445 (<0.001)

that all four consumption time series (total real consumption, real durables consumption, real nondurables consumption and real services consumption) are Granger caused by the consumer sentiment index in the long run. However, concerning short run Granger causality, only services consumption seems to be Granger caused by the consumer sentiment index,

with a  $p$ -value of 0.062. Short run changes in the consumers' confidence do not seem to contain much information on future consumption of durables or nondurables. The level of the CSI contained in the long run cointegration relation has a much more pronounced impact.

## 4.2 Time decomposition

Say we want to forecast the consumption of next month. Then, the measures of causality at lag  $j$ ,  $C_j$ , tell us what the relevant information is in the consumer sentiment index of  $j$  months ago. These measures are essentially comparing the models (2.12) and (2.13). The interpretation of  $C_j$  is then as follows: it indicates how important the CSI measured at  $t - j$  is to forecast consumption at  $t$ , given that the outcomes of the CSI are already disposable for periods  $t - j - 1$ ,  $t - j - 2$ ,  $\dots$ . In particular, if all  $C_j$ 's are zero for  $j > j_0$ , then it is sufficient to use only the last  $j_0$  indices to make the forecast. On the other hand, if all  $C_j$ 's are non significant for  $j < j_0$ , then the most recent surveys are not adding much information. If the latter would be the case, one could argue that it is too expensive to carry out monthly surveys, and decrease the frequency for collecting the survey data.

Figure 1 plots the measures of causality of the consumer sentiment at lag  $j$  for total real consumption and real consumption of durables, nondurables and services respectively. Each value of  $C_j$  can be compared with the critical value at the 5% level, indicated by the horizontal line. The maximum lag we considered was  $M = 18$ . Several checks confirm that the residuals with  $M = 18$  are white noise and do not contain any forecasting information. Moreover, it seems indeed reasonable, and it is confirmed by the graphs in Figure 1, that the consumer sentiment of more than a year and a half ago does not contain much information about today's consumption. As such, a graphical representation of the decomposition of the predictive power of the consumer sentiment index is obtained.

Figure 1 shows that, as  $j$  grows, the  $C_j$ 's tend to zero. This means that by going far enough back in time, we will reach a point where the lagged consumer sentiment index does not contain any additional information about today's consumption. For total real consumption (Figure 1a) we see that  $C_1$  is by far the most important, followed by smaller Granger causality measures at the higher order lags. This implies that, to forecast the total real consumption

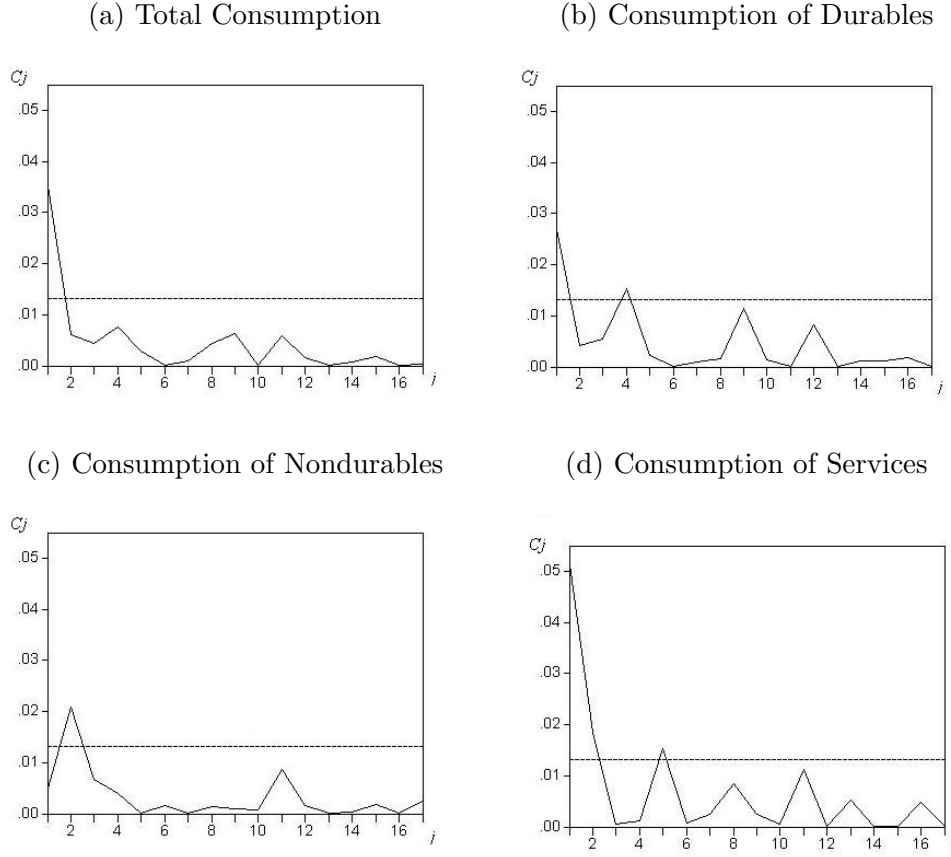


Figure 1: Time decomposition of the Granger causality of CSI for (a) total consumption, (b) consumption of durables (c) consumption of nondurables and (d) consumption of services. The dotted line gives the 5% critical value.

in the next month, using the CSI of this month is most essential. In particular, carrying out monthly surveys is worth doing. For the 3 components of consumptions, the picture is slightly different.

First of all, for durables (Figure 1b) there are peaks at different lags and the relative contribution of the higher order lags to the total measure of Granger Causality  $C$  is more important. To forecast the consumption of durables in the next month, not only this month's consumer sentiment index is informative, but also those up to at least one quarter ago. This result is in line with our primary belief that the consumption of often expensive durables is easily postponed. For services (Figure 1d) a similar picture arises, but there the contribution

of the most recent CSI to  $C$  is much more important. A possible explication for this is that “services” is a very heterogeneous consumption category, including items like recreation (where the willingness to buy will quickly be followed by actual consumption) and renting houses (being more comparable to a durable good).

For the nondurables (Figure 1c), only the Granger causality measure at lag two is significant. It turns out that the most recent value of CSI is adding much less here than for services and durables. This is at first sight surprising, but it can be explained by the fact that consumption of non-durables is already quite easy to predict from its own past. People have to fulfil their primary needs such as food, and will not change too abruptly their buying behavior for nondurables. For these goods, it is much harder for the CSI to add predictive power. This is in contrast with services, which are more difficult to predict from their own past, since their consumption is more subject to sentiments and impulsive behavior. Note that services make up about 59% of total personal consumption expenditure, compared to 12% for durables and 29% for nondurables.

Table 2 presents the mean causality lag  $D$ , together with the total causality measure for the four consumption series. We see that the consumption series are led by the consumer sentiment index with an average time lag between 4 to 6 months. While most predictive content is in the last survey value of the CSI, lagged values cannot be neglected and need to be taken along. The total Granger causality measure  $D$  is highest for the consumption of services, followed by durables and nondurables. As consumers feel more or less confident, they adapt more easily their consumption level in services (recreation, restaurants, bank services, ...) than in goods. The higher total effect on durable goods compared to nondurables is in line with the findings of Throop (1992) and with the liquidity hypothesis of Mishkin (1976) which states that consumers who fear financial distress in the near future, and thus having low confidence, will prefer to hold their wealth in liquid form. Spending on nondurables can more difficultly be delayed in time as they are usually of primary need. These findings are in line with Delorme, Kamerschen and Voeks (2001) who also found that in the USA the confidence index has more predictive power for durables than for non-durables.

Table 2: Mean lag and measure of the Granger causality of US consumer sentiment for real US consumption.

	total consumption	durables consumption	nondurables consumption	service consumption
mean lag $D$	4.2	5.2	5.5	4.4
measure of causality $C$	0.0783	0.0822	0.0562	0.123

### 4.3 Forecasting power at different horizons

So far, we have only looked at one-step ahead forecasts of the consumption series. To gain insight on how consumer confidence helps to forecast consumption at larger time horizons, we will compare the  $R^2$  of the following two equations:

$$\Delta Y_{t+h} = \alpha + \sum_{k=0}^M \beta_k \Delta Y_{t-k} + \varepsilon_{t+h}, \quad (4.1)$$

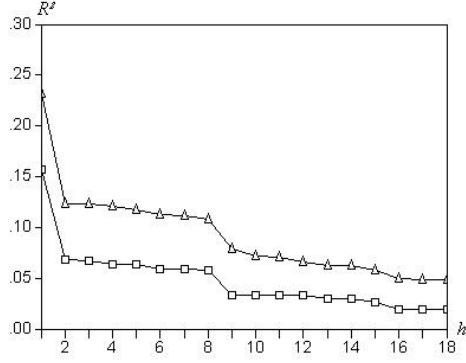
and

$$\Delta Y_{t+h} = \alpha + \sum_{k=0}^M \beta_k \Delta Y_{t-k} + \sum_{k=0}^M \gamma_k \Delta X_{t-k} + \delta U_t + \varepsilon_{t+h}, \quad (4.2)$$

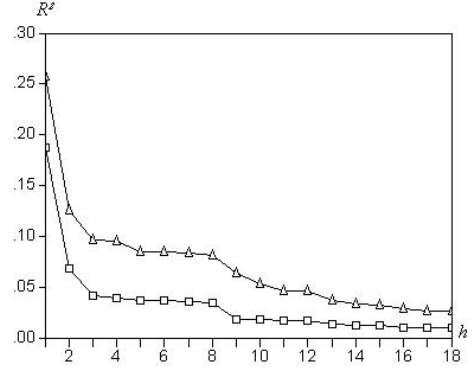
where  $Y_t$  equals consecutively the logarithm of each of the four consumption series,  $X_t$  is the consumer sentiment index and  $U_t$  equals the error correction term. Figure 2 gives the evolution of the  $R^2$  of equations (4.1) and (4.2) as a function of the forecast horizons  $h$  for the four consumption series. Obviously, as can be seen from Figure 2, the predictive power decreases in  $h$ . The difference between the two lines in the graphs is of main interest here. The difference between the  $R^2$  of equation (4.2) and the  $R^2$  of equation (4.1) reveals how much is additionally explained by the present and past values of the consumer sentiment index. This difference can be interpreted as a measure for forecasting power of the CSI at horizon  $h$ .

The results for  $h$  equal to 1, the one-step ahead forecast, are presented in Table 3. In a previous study, Carroll, Fuhrer and Wilcox (1994) report that the lagged consumer sentiment index on its own explains about 14% in the variation of the growth of total personal expenditures. However, they did not control for the past values of consumption and neither did they include the long run relationship. Making a forecast with the simple univariate model

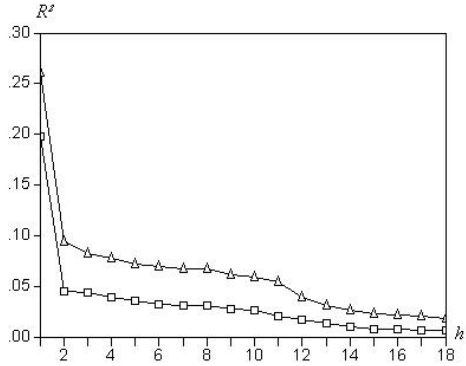
(a) Total Consumption



(b) Consumption of Durables



(c) Consumption of Nondurables



(d) Consumption of Services

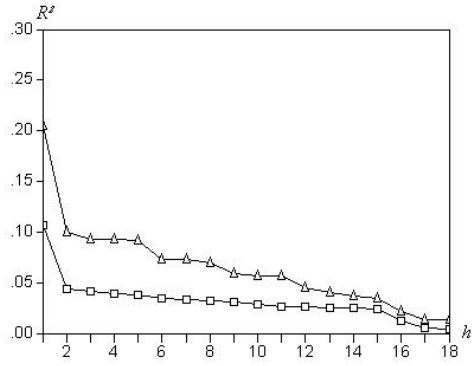


Figure 2: The  $R^2$  measure of the forecasting model using only the past of consumption ( $\square$ -line) and the model containing both the past of consumer sentiment and the past of consumption ( $\triangle$ -line) versus the forecast horizon for (a) total consumption, (b) consumption of durables (c) consumption of nondurables and (d) consumption of services.



Table 3: The  $R^2$  measure for a one step ahead forecast using the past of consumption only and using both the past of consumption and the past of consumer sentiment

consumption $Y_t$	Explained by	Additionally explained	Total
	the past of $Y_t$	by the past of $CSI_t$	
total	15.7	8.2	23.9
durables	18.7	7.6	26.3
nondurables	19.8	5.8	25.6
services	10.7	10.8	21.5

(4.1) yields indeed an  $R^2$  measure of around 15%, as can be seen from Table 3. Adding the consumer sentiment index as a predictor in a VEC model, increases the explained variance further by 5 to 10 percent. So the consumer sentiment index does serve as a valuable predictor and gives timely information of future consumption. Comparing the  $R^2$  measures for the different components of consumption reveals that the consumer sentiment adds most additional information for future consumption of services, in comparison to the other consumption series (being consistent with the total measures of Granger causality in Table 2). A reason for this may be that consumption of services is more depending on the willingness to buy, than on the ability to buy. On the other hand, consumption of durables may depend more on ability than on willingness to buy, as was also found by Van Raaij and Gianotten (1990). The additional predictive power of the CSI is lowest for non-durables. The reason for this, as we can infer from Table 3, is that this series is more easily predictable from its own past: consumers tend to have their own buying patterns for many non-durable consumption goods, which they are not changing too much on the very short run. The total  $R^2$  for non-durables is, however, still fairly high.

Figure 2 gives additional insight in the predictive power of the CSI at larger time horizons. First notice the sharp drop in  $R^2$  after the one-step ahead forecast, especially for nondurables. This drop in total predictive power is mainly due to a decrease of the part of the  $R^2$  that can be attributed to past consumption only. The extra gain in  $R^2$  given by the consumer sentiment index is, however, quite persistent. This shows that the CSI is indeed able to pick

up a latent sentiment of the consumer, which will be reflected in his buying pattern over a longer time horizon. We conclude that the monthly surveys do have a longer term predictive content, for time horizons up to even one year.

## 5 Conclusion

This paper investigates whether the US consumer sentiment index offers timely information about future US consumer spending. Our contribution is threefold. First, while most previous analyses conduct a Vector Autoregression model, we find consumption and the consumer sentiment index to be cointegrated and therefore we test for Granger causality in a Vector Error Correction framework. This allows us to take into account the long run relationship between the two series. By doing so, a very strong long run Granger causality relation was found.

Secondly, we computed total Granger Causality measures and decomposed it over different time lags. A graphical representation of this decomposition, as in Figure 1, turns out to be useful. We find that for predicting the total consumption of next month, today's sentiment index is in general most informative, but further lags need to be taken into account as well, as is confirmed by an average causality lag between 4 and 6 months. Regarding the different components of Granger Causality, we find that the CSI has most predictive content for future spending on services. Somehow surprisingly, it turned out that for nondurables, the incremental predictive power of CSI is more limited than for durables and services.

Thirdly, we measured the forecasting power of the consumer sentiment index at different forecasting horizons. We conclude that the consumer sentiment remains a useful predictor of consumption for larger time horizons.

A Granger causality analysis is most often represented by just one number: the outcome of the test statistic. In this paper, a more complete analysis is carried out, by decomposing the causality in the time domain applied to consumer sentiment and consumer spending. It became clear that such an approach yields new and interesting additional insights in the causal relationship.

The methodology outlined in this paper can be applied to other settings where the forecasting performance of a leading indicator needs to be analyzed. In this paper the consumer

sentiment index is taken as a leading indicator for consumption, but there are other possible indicators. In a study of Krystalogianni Matysiak and Tsolacos (2004) more than 20 possible economic indicators for commercial real estate performance in the UK are investigated. Other examples are Lemmens, Croux and Dekimpe (2005), where the Granger causality of Business Tendency Surveys for national accounts series has been analyzed, and Binner, Bissoondeal and Mullineux (2005) where indicators for turning points in inflation rate are studied.

**Acknowledgements:** This research has been supported by the Research Fund K.U. Leuven and the Fonds voor Wetenschappelijk Onderzoek (Contract number G.0594.05).

## References

- Batchelor, R. and Dua, P. (1998) Improving macro-economic forecasts, the role of consumer confidence, *International Journal of Forecasting*, **14**, 71–81.
- Binner, J.M., Bissoondeal, R.K. and Mullineux, A.W. (2005) A composite leading indicator of the inflation cycle for the Euro area, *Applied Economics*, **37**, 1257–1266.
- Carroll, C.D., Fuhrer, J.C. and Wilcox, D.W. (1994) Does consumer sentiment forecast household spending? If so, why? *American Economic Review*, **84**, 1397–1408.
- Delorme, C.D., Kamerschen, D.R. and Voeks, L.F. (2001) Consumer confidence and rational expectations in the United States compared with the United Kingdom, *Applied Economics*, **33**, 863–869.
- Desroches, B. and Gosselin, M.A. (2002) The usefulness of consumer confidence indexes in the United States, *Bank of Canada working paper*, 2002-22.
- Easaw, J.Z. and 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.
- Engle, R.F. and Granger, C.W.J. (1987) Cointegration and error correction - representation, estimation, and testing, *Econometrica*, **55**, 251–276.
- Eppright, D.R., Arguea, N.M. and Huth, W.L. (1998) Aggregate consumer expectation indexes as indicators of future consumer expenditures, *Journal of Economic Psychology*, **19**, 215–235.
- Gouriéroux, C., Monfort, A. and Renault, E. (1987) Kullback causality measures, *Annales D'Economie et de Statistique*, **6/7**, 369–410.

- Gouriéroux, C. and Monfort, A. (1990) Time Series and Dynamic Models, Cambridge University Press, Glasgow.
- Gouriéroux, C. and Jasiak, J. (2001) Financial Econometrics, Princeton University Press, Oxford.
- Huth, W.L., Eppright, D.R. and Taube, P.M. (1994) The indexes of consumer sentiment and confidence - leading or misleading guides to future buyer behavior, *Journal of Business Research*, **29**, 199–206.
- Johansen, S. (1995) Likelihood-based inference with cointegrated vector-autoregressive models, Oxford University Press, New York.
- Krystalogianni, A., Matysiak, G. and Tsolacos, S. (2004) Forecasting UK commercial real estate cycle phases with leading indicators: a probit approach, *Applied Economics*, **36**, 2347–2356.
- Kumar, V., Leone, R.P. and Gaskins, J.N. (1995) Aggregate and disaggregate sector forecasting using consumer confidence measures, *International Journal of Forecasting*, **11**, 361–377.
- Lemmens, A., Croux, C. and Dekimpe, M. (2005) On the predictive content of production surveys: a pan-European study, *International Journal of Forecasting*, **2**, 363–367.
- Ludvigson, S.C. (2004) Consumer confidence and consumer spending, *Journal of Economic Perspectives*, **18**, 29–50.
- Mishkin, F.S. (1976) Illiquidity, consumer durable expenditure, and monetary policy, *American Economic Review*, **66**, 642–654.
- Miller, S.M. and Russek, F.S. (1990) Cointegration and error-correction models: the temporal causality between government taxes and spending, *Southern Economic Journal*, **57**, 221–229.
- Roberts, I. and Simon, J. (2001) What do sentiment surveys measure? *Research Discussion Paper. Economic Research Reserve Bank of Australia*.
- Souleles, N.S. (2004) Expectations, heterogeneous forecast errors, and consumption: micro evidence from the Michigan Consumer Sentiment Surveys, *Journal of Money, Credit and Banking*, **36**, 39–72.
- Throop, A.W. (1992) Consumer sentiment: its causes and effects, *Economic Review, Federal Reserve Bank of San Francisco*, 1992-nr1.
- Toda, H.Y. and Phillips, P.C.B. (1994) Vector Autoregression and Causality: A Theoretical Overview and Simulation Study, *Econometric Reviews*, **13**, 259–285.

- Utaka, A. (2003) Confidence and the real economy - the Japanese case, *Applied Economics*, **35**, 337–342.
- Van Raaij, W.F. and Gianotten, H.J. (1990) Consumer confidence, expenditure, saving, and credit, *Journal of Economic Psychology*, **11**, 269–290.
- Vuchelen, J. (2004) Consumer sentiment and macroeconomic forecasts, *Journal of Economic Psychology*, **25**, 493–506.