Consumer confidence and consumption forecast: a non-parametric approach

Bruno, Giancarlo

ISTAT - Rome (Italy)

September 2012

Consumer confidence and consumption forecast: a non-parametric approach

Giancarlo Bruno*

Istat

The consumer confidence index is a highly observed indicator among short-term analysts and news reporters and it is generally considered to convey some useful information about the short-term evolution of consumer expenditure.

Notwithstanding this, its usefulness in forecasting aggregate consumption is sometimes questioned in empirical studies. Overall, the conclusions seem to be that the extensive press coverage about this indicator is somewhat undue.

Nevertheless, from time to time, attention revamps on consumer confidence, especially when turns of the business cycle are expected and/or abrupt changes in this indicator occur. Some authors argue that in such events consumer confidence is a more relevant variable in predicting consumption. This fact can be a signal that a linear functional form is inadequate to explain the relationship between these two variables. Nevertheless, the choice of a suitable non-linear model is not straightforward.

Here I propose that a non-parametric model can be a possibile choice, in order to explore the usefulness of confidence in forecasting consumption, without making too restrictive assumptions about the functional form to use.

1. Introduction

In this paper the relationship between aggregate consumer expenditure and the consumer confidence index for Italy is enquired. In particular, the focus concerns the usefulness

*E-mail: `giancarlo.bruno@istat.it`

of the latter variable to get better forecasts of consumer expenditure. This subject, although it has been explored in many other papers, in my opinion still deserves some examination.

The consumer confidence index is usually built as a composite indicator by averaging the answers of a sample of consumers over a range of questions possibly linked to the economic situation as they live and perceive it. This indicator receives considerable attention on the daily press as it is seen as an early available variable useful to predict aggregate consumption. In particular, it is supposed to mirror aspects of consumers' sentiment which can have an impact on their propensity to consume (Roos, 2008).

From a forecaster point of view, early availability and absence of revisions of this indicator are key aspects to focus the attention on; indeed, quarterly national accounts data on consumption are first released after 70 days following the end of the period they refer to and are subject to deep revisions. Moreover, they are available only at quarterly frequency, while confidence is recorded monthly. On the other hand, empirical results on effective forecasting ability of consumer confidence index are mixed, and probably below the expectations implied by the daily press attention it receives.

Here, I argue that a possible weakness in many estimated empirical models can be attributed to the use of a linear functional form to model the relationship between consumption and confidence. The use of linear models, in fact, can be partially inconsistent with the view that only abrupt shifts in consumer confidence are relevant to signal changes in consumer expenditure. In addition to this effect, which can be denoted as a threshold effect, asymmetries over the business cycle could also characterize such a relationship. Both these features could be better picked-up with the use of a non-linear model. Nevertheless, the specification of a suitable non-linear model is not that easy. In this paper, in order to facilitate the problem a non-parametric approach is utilised.

The paper is organised as follows. Main results in the literature concerning consumption and confidence are reported in the next section. Section 3 describes the data used, section 4 illustrates the model. Identification and estimation are dealt with in section 5, where it is given also an illustrative example of the use of the proposed model as a descriptive device. The results of the forecasting comparison are displayed in section 6 and a final section concludes.

2. Literature and motivation

The possible relevance of consumer confidence in helping forecasting consumption has a long history which paralleled the theoretical analysis on consumer behaviour.

The stream of studies in psychological analysis of consumer behaviour underline the importance of consumers' moods and attitudes as independent causes of consumer spending together with income. This would be true, in particular, for what are defined as discretionary expenses, which are generally identified with durables (Roos, 2008). On the other hand, the most accepted theory of consumer behaviour, the life-cycle theory/permanent income hypothesis (PIH), in its strict form does not allow for a predictive power of the confidence index over consumption, beyond its capacity to signal changes in permanent income (McIntyre, 2007).

Actually, a predictive power of confidence on consumption is often found: for example, Fuhrer (1988), Carroll et al. (1994) and Bram and Ludvigson (1998) all find that including confidence improves short-term consumption forecasts. Attempts have been made, therefore, to include this indicator in the theoretical framework of the PIH as a further explanatory variable of consumer behaviour. McIntyre (2007) develops a "confidence augmented" permanent income model: the empirical analysis conducted there suggests that including consumer confidence in the model substantially decreases estimates of the intertemporal elasticity of substitution. Also Roos (2008) tries to include confidence factors in a standard model of intertemporal utility maximization.

On the other hand, results on forecasting usefulness of confidence index over consumption expenditure are not always very sharp. For example, in the cited study of Bram and Ludvigson (1998), results are seldom significant, according to the predictive accuracy test of Diebold and Mariano (1995). In addition, in a more recent contribution Ludvigson (2004) supports the view that there is relatively modest independent information coming from the consumer confidence index not already included in other variables. Also, according to Al-Eyd et al. (2009) the relation between consumption and confidence would be unstable over time. The authors analyse European data and find very limited predictive content for the case of Italy, both in a relation between consumption and confidence, as well as considering other variables. A cautious view on the use of consumer confidence for forecasting is also contained in Batchelor and Dua (1998).

A related issue is the role of business cycle in the problem at hand. Indeed, a possible reason behind the usefulness of confidence indicators is that they can signal early a business cycle turning point. In this respect Matsusaka and Sbordone (1995) find that consumer confidence accounts for about 20% of business cycle innovation in U.S. GDP. Moreover, Acemoglu and Scott (1994) note that UK consumer confidence predicts consumption even after conditioning on income. Their suggestion to explain this fact is that the consumption function shifts over the business cycle.

Indeed, the relation between confidence and consumption cannot be considered irrespectively of the purpose of the analysis. The objective, in fact, can be either explaining consumer behaviour or a pure forecasting issue. In this contribution I am focusing exclusively on the second aspect. Carroll et al. (1994) are particularly clear on the key issues for forecasting. With respect to this point they identify two questions: 1) if the consumer confidence has predictive content for future changes in spending; 2) if this information adds something to that contained in other indicators. To these points I would add the fundamental issue of actual data availability at a given point of time: if we consider this, usually only confidence indexes and some financial variables are available almost in real time. Therefore, even though confidence simply summarizes information contained in other variables, such as income etc., without adding anything new, it can be nonetheless useful in a forecasting model as long as the other variables are not yet available at the moment of forecast production.

In this paper, I point out that the extraction of a predictive content from the confidence indicators can be significantly improved, concentrating on different aspects. First of all, the link between aggregate consumption and confidence indicator is usually taken as linear. However, this assumption has no firm theoretical foundation, apart from its more practical implementation. In a recent study Qiao et al. (2009) propose a non-linear causality test, and conclude advocating the use of non-linear forecasting models for expressing the relation between confidence and consumption.

Here I share this view; nevertheless, this raises the question of which model among the many non-linear ones available is better suited to this purpose; in fact, there is no clear guidance as to what should be the better choice for a non-linear model. Some works on this point suggest that a threshold effect could be present (Desroches and Gosselin, 2002). I partly overcome this problem leaving the functional form linking consumption and confidence unspecified and estimating it with a non-parametric approach.

Secondly, most of the times consumer confidence is taken "as is", whereas a different weighting scheme of the different components could in principle improve its predictive ability. In this paper I also use an alternative weighting scheme, where the weights are chosen in a framework consistent with the non-linear model proposed.

In the end, as well as aggregate private consumption, I take into account also consumption broken down by durability. This is consistent with the theoretical works which stress the different influence of confidence on durables with respect to non-durables goods. In particular Katona (1971) supports the view that "movements of the index of consumer sentiment [...] serve to indicate changes in the direction of prospective discretionary expenditures". Durables have often been identified as natural candidates to the role of goods having discretionary nature, even though nowadays this kind of distinction is probably less clear cut.

3. Data

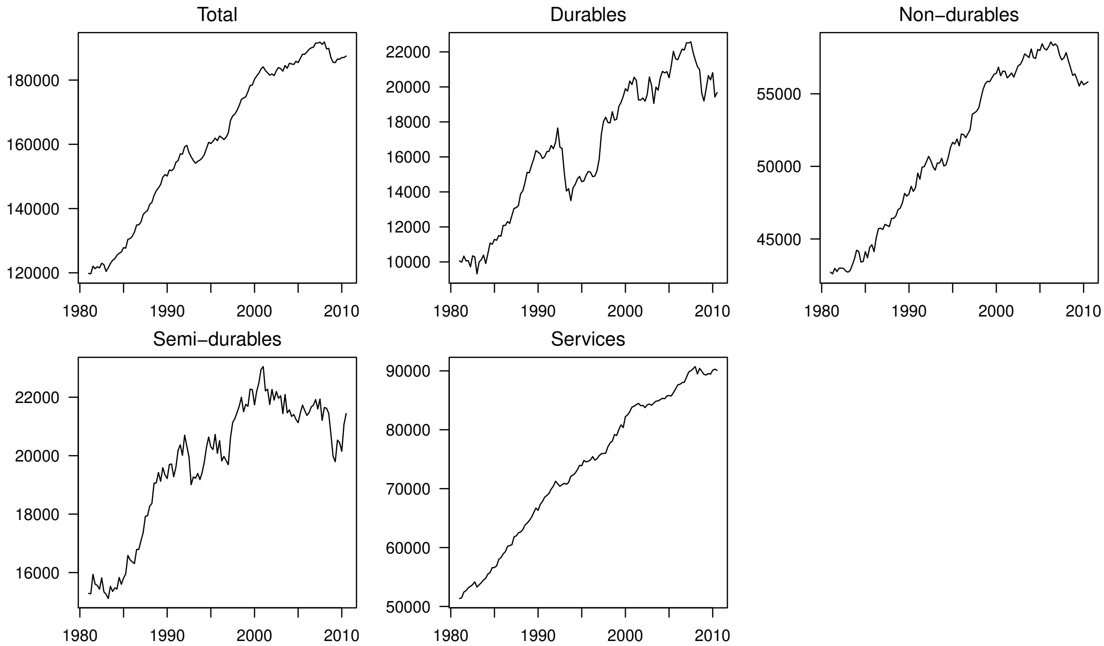
The data used in this paper refer to Italy and come from two main sources: consumer expenditure data produced in the framework of quarterly national accounts and a consumer confidence indicator. The common span available is from first quarter 1982 to third quarter 2010.

Consumer expenditure data are recorded at quarterly frequency, seasonally adjusted, broken down by durability: durables, non-durables, semi-durables, services. The reason why this kind of breakdown has been chosen is that the traditional literature in this field underlines that the impact of consumer confidence could be more relevant in the case of durable goods. The consumption data are volume data, obtained with a chained Laspeyres formula so as to minimize the distortionary effect of a fixed weighting scheme. The data are partly calculated interpolating annual figures by means of related quarterly indicators, following the procedure described in Chow and Lin (1971). Quarterly consumption data are first published 70 days after the end of the reference period and subsequently revised every quarter for at least three years. Moreover, sometimes they undergo major revisions due to exceptional causes like methodological modifications, classification changes, etc. All the consumption series appear to contain a unit root, therefore I take as the stationary transformation the first difference of logs.

Consumer confidence is produced starting from the results of a monthly survey carried out by Istat[1] on a random sample of 2000 households. Households are asked a number of questions about their present and future personal situation as well as how they judge and forecast some macro variables. Some of the questions, namely those composing the confidence index, are listed in appendix A. Every question can be given a qualitative ordinal answer; most of the times there are five options, ranging from “a lot better” to “a lot worse”. Aggregating over individuals one obtains relative frequencies of the different answers. A further step consists in synthesizing the answers for every given question (so called *quantification*). There are many methods proposed in the literature to cope with this issue (for an extensive review see Proietti and Frale, 2007). A simple and frequently used method is the so called *balance*, where the share of negative answers (“a lot worse” and “worse”) are subtracted from the positive ones (“a lot better” and “better”). Here we adopt the method currently implemented for the published consumer

[1] The survey has been carried out for several years by ISAE - *Istituto di studi e analisi economica*- a public owned economic research institute suppressed in 2010 and whose functions were partly attributed to Istat.

Figure 1: Consumption series by durability - levels



confidence index, which uses a balance where the extremes are given twice the weight of the other answers.

Balance of question i at time t is therefore given by:

$$x_{it} = 2{\cdot}BB_{it} + B_{it} - W_{it} - 2{\cdot}WW_{it} \tag{1}$$

where BB_{it}, B_{it}, W_{it}, and WW_{it} are, respectively, the percentage shares of answers "a lot better", "better", "worse" and "a lot worse" for question i in period t. The quantified series, therefore, is a bounded one whose range is in the interval $[-200, 200]$.

In the following step a simple average of the nine balances so obtained is taken,[2] resulting in the confidence indicator (CF):

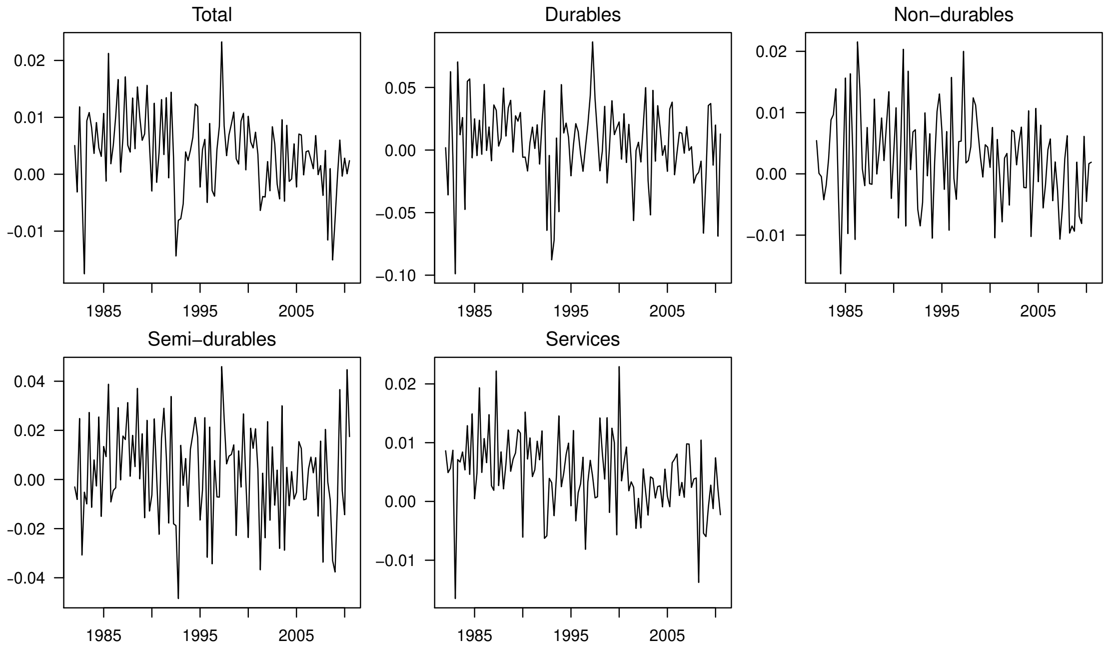
$$CF_t = \frac{1}{9}\sum\nolimits_{i=1}^{9} x_{it} \quad . \tag{2}$$

The confidence index is usually scaled to be 100 in a certain year. Here I simply scale it, with no loss of generality, in order to have a bounded variable in the interval $[0, 1]$.

The confidence indicator is recorded monthly, while consumption data are available

[2]The series of unemployment prospects enters the confidence indicator with inverted sign.

Figure 2: Consumption series by durability - first difference of logs



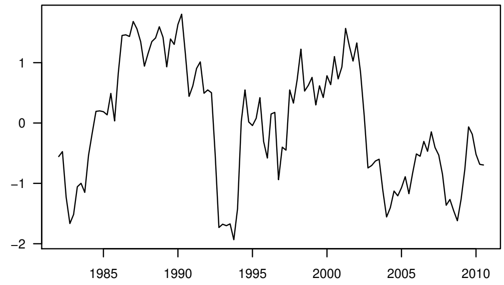
only at quarterly frequency. The forecasting model is then built at quarterly frequency aggregating over time the confidence indicator with a simple average. The use of a simple average rather than other schemes (e.g. the value of the last month of the quarter) has the advantage of smoothing out very short-term variability of the series.

The very early availability of this indicator makes it particularly attractive for short-term forecasting. In fact, at the end of month t the indicator referring to that month is immediately available. If we take the end of a quarter as our forecasting production period, we have data on consumption available up to the previous quarter and confidence data up to the quarter itself. Therefore, even in the absence of a leading nature of the confidence index, it can be useful in a prediction context, due to its earlier availability.

4. Model formulation

The information set considered therefore is composed by the consumption aggregate (each time, in turn, total, durables, non-durables, semi-durables, services) and the confidence indicator. Other possible explanatory variables are excluded to be consistent with the ultimate goal of this exercise, that is, short-term forecasting. In fact, some important variables deemed to explain aggregate consumption, such as disposable income,

Figure 3: Consumer confidence index



wealth, ecc. are available, if at all, many months after the end of the reference period, so I exclude them from this exercise. As stated before, if the confidence indicator has a coincident nature with respect to consumption, one step-ahead forecasts of the latter can be generated given the publication schedule of these variables. Moreover, if the confidence indicator has a lead of one quarter, then consumption forecasts are possible up to two-step ahead. Given the results of the literature on consumption and confidence this is the plausible forecasting horizon we can test with such an information set.

Let us denote the consumption at quarter t by C_t (this actually indicates, in turn, total consumption, durables, semi-durables, non-durables, services) and the confidence indicator by CF_t. The non-linear non-parametric model I propose to use is a functional coefficient regression (FCR) model.

$$\Delta \log(C_t) = f_0(CF_t) + \sum_{i=1}^{p} f_i(CF_t) \Delta \log(C_{t-i}) + \varepsilon_t. \tag{3}$$

The main features of model (3) are summarized in section 5, together with estimation and inference issues. The coefficients f_i are not specified within a fixed parametric functional form: rather, their form is estimated from the data with a non parametric technique, namely a local linear smoother. The FCR model nests as particular cases the linear AR model, as well as TAR and EXPAR models.

The proposed formulation allows to accomodate a non-linear relation between consumer expenditure and consumer confidence, without specifying precisely its form. However, the model is less prone than other more general formulations to the so called *curse of dimensionality*, since the functional coefficients $f_i(\cdot)$ are one-dimensional, therefore the estimation problem is kept feasible.

The formulation of model (3) is consistent with a coincident nature of the confidence index with respect to consumption, since it enters the model contemporaneously. Therefore, given the publication schedule, it allows to produce only one-step-ahead forecasts. In order to exploit the possible leading nature of the confidence index, a second model is also estimated (FCR2) where confidence index is lagged one period. In this way the model allows to produce up to two-step ahead forecasts:

$$\Delta \log(C_t) = f_0(CF_{t-1}) + \sum\nolimits_{i=1}^{p} f_i(CF_{t-1})\Delta \log(C_{t-i}) + \varepsilon_t. \tag{4}$$

In order to evaluate the usefulness of the proposed models a couple af bivariate linear models are also estimated, In particular, since the results of Granger causality tests exclude any feedback effect of consumption on confidence, a standard auto-regressive distributed lag (ADL) model is considered here. In particular, starting from the general formulation:

$$\Delta \log(C_t) = \alpha + \sum\nolimits_{i=1}^{4} \beta_i \Delta \log(C_{t-i}) + \sum\nolimits_{i=0}^{4} \gamma_i CF_{t-i} + \varepsilon_t. \tag{5}$$

a subset lag structure is chosen by means of the BIC criterion.

A second model ADL2 is also proposed which, analogously to what is done with FCR2, allows to obtain two-step ahead forecasts, exploiting the possible leading nature of the confidence indicator:

$$\Delta \log(C_t) = \alpha + \sum\nolimits_{i=1}^{4} \beta_i \Delta \log(C_{t-i}) + \sum\nolimits_{i=1}^{4} \gamma_i CF_{t-i} + \varepsilon_t. \tag{6}$$

The difference with model (5) is that the contemporaneous confidence indicator is not permitted here. Also in this case a subset of significant regressors is selected.

In addition, a non-linear parametric model, namely a threshold autoregressive model (TAR), is considered. Such a model can be nested in the FCR model and the main difference is that there are a discrete number of states associated to the level of the state variable. Here a model with two states is used:

$$\begin{aligned} \Delta \log(C)_t &= \varphi_{10} + \varphi_{11}\Delta \log(C)_{t-1} + \cdots + \varphi_{1p}\Delta \log(C)_{t-p} + \varepsilon_1 t \quad \text{if } CF_{t-d} < r \\ \Delta \log(C)_t &= \varphi_{20} + \varphi_{21}\Delta \log(C)_{t-1} + \cdots + \varphi_{2q}\Delta \log(C)_{t-p} + \varepsilon_2 t \quad \text{if } CF_{t-d} \geq r \end{aligned} \tag{7}$$

where d is the delay parameter and r is the threshold of the state variable which governs the changes in the regime. When $d = 0$ I denote the model as TAR and it provides one-step ahead forecasts. In the case $d = 1$ it is denoted as TAR2 and can produce predictions up to two-step ahead. The autoregressive lag order p is identified by means of the AIC criterion.

4.1. Alternative confidence indicator

Model (3) can be further generalised to allow the state variable to be a more elaborate expression. In particular, it can be a linear combination of variables, as illustrated in Fan et al. (2003). A reason to implement this kind of model could be the interest in evaluating a different aggregation scheme for the variables composing the confidence indicator, such as: $CF2_t = \sum_{i=j}^{9} \omega_j x_{jt}$. (I recall that x_{jt}, $j = 1, \ldots, 9$ are balances of answers to the questions making up the confidence index).

Therefore, the following model (FCRN) is built:

$$\Delta \log(C_t) = f_0 \left(\sum_{j=1}^{9} \omega_j x_{jt} \right) + \sum_{i=1}^{p} f_i \left(\sum_{j=1}^{9} \omega_j x_{jt} \right) \Delta \log(C_{t-i}) + \varepsilon_t. \tag{8}$$

In this model the state variable CF_t, which is a simple average of the different balances, is replaced with a weighted average. Moreover, the weights ω_j are estimated from the data, consistently with the proposed model. Both Xia and Li (1999a) and Fan et al. (2003) propose methods for estimating both the weights in the linear combination constituting the state variable ω_j and the functional coefficients f_i in a model like (8), under dependence.

Analogously to what is showed in the previous section a model with lagged confidence can be build, denoted here as FCRN2, which allows up to two-step ahead forecasts to be produced.

$$\Delta \log(C_t) = f_0 \left(\sum_{i=0}^{9} \delta_i x_{i,t-1} \right) + \sum_{i=1}^{p} f_i \left(\sum_{i=0}^{9} \delta_i x_{i,t-1} \right) \Delta \log(C_{t-i}) + \varepsilon_t. \tag{9}$$

5. Identification and Estimation

As stated earlier in this paper I adopt a non-parametric approach, which means in practice that no strict hypothesis is made about the form of the functions $f_i(\cdot)$ in equations (3), (4), (8) and (9), apart from some smoothness conditions such continuity and second order differentiability.

In order to estimate these functions a local polynomial approach can be used: here, in particular, I use a local linear smoother. In the case of equation (3), considering a generic point u and letting $U_t = CF_{t-d}$ and $X_t = \Delta \log(C_t)$ this amounts to minimize the following function:

$$\sum_{t=p+1}^{T} \left\{ X_t - a_0 - b_0(U_t - u) - \sum_{i=1}^{p} \left[a_i + b_i(U_t - u) \right] X_{t-i} \right\}^2 \frac{1}{h} K \left(\frac{U_t - u}{h} \right) \tag{10}$$

with respect to $\{a_i, b_i, i = 0, \ldots, p)\}$. The local linear estimates of the functional coefficients are then simply the values $\hat{a}_i(u)$.

Equation (10) depends on some values which must be established and/or estimated from the data, namely the order p, the value d, the kernel function $K(\cdot)$ and the bandwidth h.

As regards the value of d, based on the schedule of data availability I use both the value $d = 0$, corresponding to model (3), which allows to get one-step ahead forecasts, and $d = 1$, corresponding to model (4), which implies a leading nature of the confidence index and allows to get one and two-step ahead forecasts.

For every given d, the value of p is chosen based upon a form of cross-validation similar to that proposed in Cai et al. (2000). More in detail, letting Q and m be two positive integers such that $Qm < T$, the first Q sub-series of length $T - qm$ are used $(q = 1, \ldots, Q)$ to estimate the model and then the one-step prediction errors are computed for the next m points of the series. For every value of p one can define the average prediction error:

$$APE_q(p) = \frac{1}{m} \sum\nolimits_{t=T-qm+1}^{T-qm+m} \left[X_t - \hat{a}_{0,q} - \sum\nolimits_{j=1}^{p} \hat{a}_{j,q}(U_t) X_{t-j} \right]^2 . \tag{11}$$

Moreover, averaging over the Q set of prediction errors one obtains the quantity:

$$APE(p) = \frac{1}{Q} \sum\nolimits_{q=1}^{Q} APE_q(p). \tag{12}$$

The value p is then chosen such that $APE(p)$ is minimized.

The kernel function $K(\cdot)$, which downweights observations far from the evaluation point u is chosen to be the Gaussian probability density function; this choice is probably not as determinant as the bandwidth h, which represents, roughly speaking, the width of the interval around a point u which is used for estimating the coefficient functions at that point. The bandwidth is chosen by means of the cross validation criterion. This value is maintained fixed over the support of u.

Table 1 reports the lags selected, respectively, for the autoregressive (p) and confidence lags (q) polynomials of the ADL and ADL2 models, while in table 2 are reported the lags selected for all the FCR models.

5.1. Estimation results for durable goods

Detailed comments on forecasting results are reported in the next section. Instead, here I illustrate more in detail some estimation results of the FCR model for durable goods. Similar results are available for the other sectors and models and, indeed, they are

Table 1: Lags selected for ADL models

	polynomial	total	durables	non-durables	semi-durables	services
ADL	p			(1,2,4)	(1)	(1)
	q	(0,3)	(0,3)	(0)	(0,2)	(0)
ADL2	p			(1,2,4)	(1,4)	(1)
	q	(1)		(1)	(1,2)	(1)

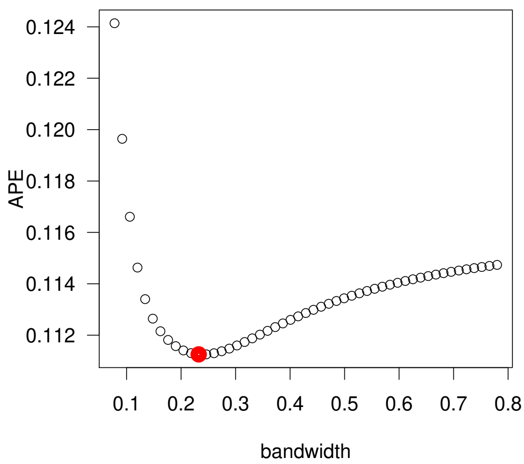
Table 2: Lags selected for FCR models

	total	durables	non-durables	semi-durables	services
FCR	4	1	1	1	1
FCRN	4	2	1	1	1
FCR2	1	1	1	1	1
FCRN2	1	1	1	4	1

qualitatively similar to those displayed here. Anyway, I focus the attention on the durable goods because in this case confidence is expected to perform better in forecasting, and, indeed, it does. Moreover, I think also that the following results can also help explaining some possible causes of the conflicting findings in the literature about the predictive content of consumer confidence.

First of all, figure 4 reports a scatter plot of the average prediction error (12) for different bandwidth values.

Figure 4: FCR model: bandwidth selection

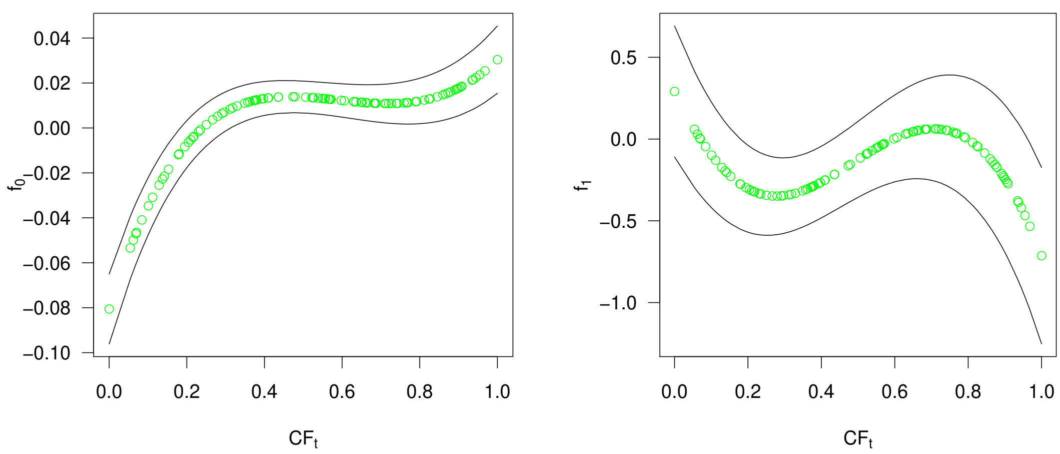


Once defined the order model p and the bandwidth h it is possible to estimate the coefficients; here this is carried out by means of a local linear smoother. Moreover, it is possible to calculate confidence bands for the functional coefficients and testing against their value being 0 or a constant or, in general, a known functional form. This is done deriving the distribution of the maximum discrepancy between estimated coefficients and the true value. A complete treatment of asymptotics under dependence for estimating FCR models is contained in Xia and Li (1999b), as well as the construction of confidence bands for the functional coefficients.

In figure 5 it is possible to look at the estimates of the two functional coefficients with 95% confidence bands. It appears from the plot that the coefficients are truly varying, and this is confirmed by the formal test. It is also interesting the possible interpretation of the coefficients, which makes this kind of model also appealing as a descriptive device. In fact, it is possible to appreciate that effectively there seems to be a sort of threshold effect: in particular it is possible to see that the coefficient f_0, which plays the role of a constant in a linear model and that in this context can be thought of as a variable drift in durable consumption, shows a sharp and sudden decline when confidence falls below 0.20. Conversely, values of confidence from 0.4 and above are associated, fundamentally, with a stabilisation of the drift. Only very high confidence values imply a slightly further increase.

The coefficient f_1, which is the first order functional coefficient, displays only a moderate variation and is concentrated around small negative values; for high values of the confidence this coefficient becomes slightly more negative.

Figure 5: FCR model: functional coefficients and confidence bands



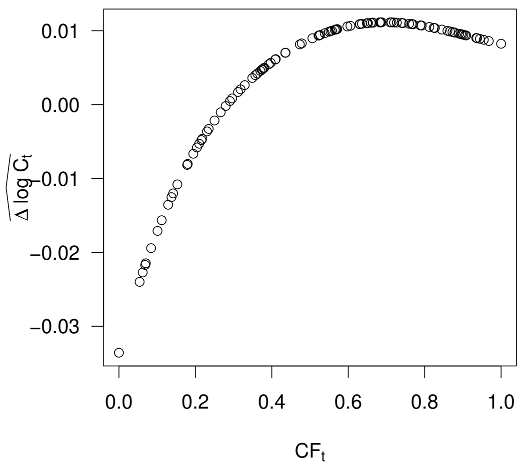
Overall the response of durables consumption is dominated by the first term, that is the variable drift. In figure 6 the overall effect is displayed for the response variable for different values of the confidence indicator. In particular I calculated the fitted values when the lagged consumption value is taken at its average value:

$$\Delta \log(\hat{C}_t) = \hat{f}_0(CF_t) + \hat{f}_1(CF_t)\Delta \log(\bar{C}) \tag{13}$$

Figure 6 illustrates that the drift coefficient f_0 dominates the response of the model. Expected consumption is always very negative when particularly low values of confidence emerge. It appears also that this effect is not at all symmetric, as very high confidence values are not exerting a significant effect. Therefore in this case a parametric non-linear model such as that used by Desroches and Gosselin (2002) would not be correct, as in that paper a symmetric threshold effect is considered.

I think that this helps also explaining some conflicting results in the literature about the usefulness of consumer confidence indicators. In fact, as long as the sample considered do not include periods with a particularly low value of the confidence index, the results could be quite different from those reported here. This is probably also the reason why the analysis of confidence has sometimes taken the form of studies of occasional events, such as the first Gulf war (Throop, 1992) or September 11^{th} attack (Garner, 2002).

Figure 6: FCR model: predicted values



6. Results

6.1. Experiment set-up

In the forecasting exercise in addition to ADL, FCR and TAR models I also use a benchmark univariate linear autoregressive (AR) model. This is necessary to assess if the confidence delivers relevant information in addition to that contained in the minimum information set constituted by the past history of the variables to be forecast. In particular, a subset AR model is selected by means of the BIC criterion from the following general model:

$$\Delta \log(C_t) = \alpha + \sum_{i=1}^{4} \varphi_i \Delta \log(C_{t-i}) + \varepsilon_t. \tag{14}$$

Considering all the models there are 9 sets of one step-ahead forecasts and 5 sets of two step-ahead forecasts to compare for every consumption aggregate. All the models have been identified using the first two thirds of observations.[3] As identification here I mean selecting the lags of the benchmark autoregressive model and of the linear ADL models; the order p of all the FCR models is also selected on this reduced sample size. Moreover, the weights of the linear combination constituting the state variable in equations (8) and (9) are also estimated only once, still within the reduced sample size.

The forecasts are generated with a recursive scheme: as long as a new observation is added, the coefficients of the linear models are re-estimated, as well as the bandwidth and the functional coefficients of the FCR models. The data set used in this work, once allowed for differences and lags creation, has a rather reduced length. In fact, 110 quarters are available. Of those, 73 are used for model identification, while the evaluation period is composed by the last 36 observations.

Two-step ahead forecasts for FCR models are always obtained with the plug-in approach, that is repeatedly substituting one-step ahead forecasts. In fact, the state variable is exogenous in this set-up, thus respecting one of the conditions needed to make the plug-in approach delivering a correct estimator of $E(X_{t+s}|X_t)$ for $s > 1$. The other condition, not met here, is that the functional coefficients must be known (Harvill and Ray, 2005). The alternative of using a bootstrap approach is not pursued here.

[3] A different scheme is followed for the TAR model whose lag structure is identified every new observation is available. This has been necessary in order to have results comparable to the other models. This effect can be due to the inadequacy of the model as well as to the instability of the non-linear relation, as pointed out in Teräsvirta et al. (2005). However, the main point here is that the FCR model has a better performance without suffering this stability problem.

6.2. Forecast comparison

The tables from 3 to 12 illustrate the forecasts evaluation. In particular, the mean error and the mean absolute error are presented for all the models and all the consumption aggregates. These statistics are always given as ratio to the benchmark values, given by the linear autoregressive model. Therefore a value less than one indicates an improvement with respect to the benchmark. In the case of the mean error there is the additional complication that it is possible to have a negative number if the mean error in the benchmark and in model under evaluation have a different sign; anyway, also in this case a value less than one (in absolute value) is to be intended as an improvement.

6.2.1. Total consumption

In the case of total consumption the use of the confidence indicator in a linear model improves the forecasts obtained over the benchmark univariate AR model: mean absolute error in the ADL and ADL2 models are about 0.92 relative to the benchmark at 1-step ahead, and 0.90 at 2-step ahead. Also the mean error shows a sizeable improvement when the confidence index is included in linear form. The gain in predictive accuracy is significant in the case of the ADL2 model, particularly for one-step ahead forecasts.

The use of FCR models further improves both the MAE and the ME: with respect to the benchmark the reduction is roughly 16-19% for the MAE and more than 70% when considering the mean error. The performance of models incorporating the alternative confidence indicators is resembling those which use the traditional one. Indeed, the estimated weights produce, in this case, an alternative confidence index which is very similar to the official one.

When looking at predictive accuracy tests, it turns out that at one-step ahead only the improvement for the basic FCR model is marginally significant (the p-value of the DM test is 0.10). Moreover, when considering two-step ahead forecasts the results are clearly better for the FCR model, both in the FCR2 and FCRN2 formulation. In this case forecast gain over the benchmark is almost 20% and statistically significant.

The use of a lagged confidence index does not modify in a sensible way the results, except for the TAR model, which in this case reaches a performance similar to FCR model, but even more significant, both for one and two-step ahead predictions.

6.2.2. Durables

With this aggregate a substantial difference emerges between the parametric and non-parametric models. The ADL model shows an improvement just for the mean error, while the MAE stays at 0.97 relative to the benchmark. In the case of ADL2 model, the identification of the subset does not include the confidence index at all, thus obtaining the same results of the simple autoregressive model. TAR models predictions are overall worse than the benchmark.

On the other hand the non-linear model shows here a better performance. The mean error is almost annihilated with respect to the baseline forecasts, while the mean absolute error is slightly more than 0.90 for most of the cases. Moreover, the predictive accuracy is also significantly better than the benchmark.

When evaluating two-step ahead forecasts the results are even more clear, as the lagged confidence is excluded from the ADL2 model specification. On the other hand the FCR2 and FCRN2 models see a reduction of 10% for the mean absolute error and over 90% for the ME; such a reduction is also highly significant, with a p-value for the DM test less than 0.05.

What is found here is consistent with earlier works on this subject which underline the importance of confidence especially for deciding the timing of durables expenses. Moreover, it helps explaining some seemingly contrasting results as in Malgarini and Margani (2007), where, in a different context, confidence does not appear to exert any role on durables, as long as their results could be attributed also to the use of a linear functional form.

6.2.3. Non durables

Also for this aggregate the results are clear-cut. Adding the confidence indicator in a linear model does not improve, rather worsens, sometimes significantly, the forecasting performance. In fact, ME and MAE are higher both for ADL and for ADL2 model with respect to the benchmark. Moreover, results of the Diebold-Mariano test show that this difference is often statistically significant. Just the opposite behaviour characterizes the non-linear models, particularly the FCR model. As a result the latter is significantly better than the simple autoregressive and ADL models. In addition, the mean error is about one tenth of the benchmark. Models adopting the new confidence indicator are somewhat in the middle: they show a better performance than the benchmark but results are seldom significant at conventional confidence levels. No relevant differences emerge as long as coincident or lagged confidence is used. The TAR model performance is somewhat in between the ADL and FCR one.

The results found for this aggregate are consistent with those obtained by Qiao et al. (2009) on US data. In that case the authors find that a Granger causality test fails in finding predictive content of consumer confidence on non-durables consumption, while their proposed non-linear test suggests the opposite: they conclude that a non-linear model is the natural candidate to use in that case.

6.2.4. Semi-durables

In this sector, which includes goods such as books and clothes, the confidence indicator is never adding something to the baseline AR model, both in the case of linear specification as well as in the non-linear and non-parametric one. This is also true in the case of two-step ahead forecasts. Moreover, in most of the cases, the inclusion of confidence worsens the mean absolute error with respect to the benchmark although not significantly. An improvement, instead, is observed as far as the mean error is considered.

6.2.5. Services

A similar pattern is observed for services when one-step ahead forecasts are analysed. However, in this case emerges a potential for confidence as long as only two step ahead forecasts are considered. In this sense it seems that confidence has a leading nature as long as services are taken into account. A limited improvement, although not statistically significant, is observed at one-step ahead forecasts only for the ADL model over the AR model. A more sizeable improvement is registered in the case of two-step ahead forecasts for all the models. In particular, the use of the alternative confidence indicator (FCRN2 model) improves considerably the results.

Anyway, when taking into consideration predictive accuracy with respect to the benchmark, only ADL2 displays a significant gain.

Table 3: Errors statistics relative to AR model - Total consumption - 1-step-ahead - Errors relative to AR model and probability of DM test of equal forecasting accuracy (H1: row better than column)

	AR	FCR	ADL	TAR	FCRN	FCR2	ADL2	TAR2	FCRN2
ME	1.00	-0.23	0.62	0.61	-0.24	-0.15	0.76	0.32	-0.13
MAE	1.00	0.83	0.92	1.01	0.84	0.87	0.93	0.84	0.84

	AR	FCR	ADL	TAR	FCRN	FCR2	ADL2	TAR2	FCRN2
AR		0.90	0.80	0.48	0.89	0.84	0.92	0.96	0.89
FCR	0.10		0.20	0.10	0.35	0.23	0.20	0.44	0.40
ADL	0.20	0.80		0.22	0.78	0.69	0.46	0.84	0.78
TAR	0.52	0.90	0.78		0.89	0.86	0.77	0.98	0.90
FCRN	0.11	0.65	0.22	0.11		0.28	0.22	0.48	0.47
FCR2	0.16	0.77	0.31	0.14	0.72		0.30	0.64	0.84
ADL2	0.08	0.80	0.54	0.23	0.78	0.70		0.88	0.78
TAR2	0.04	0.56	0.16	0.02	0.52	0.36	0.12		0.51
FCRN2	0.11	0.60	0.22	0.10	0.53	0.16	0.22	0.49	

Table 4: Errors statistics relative to AR model - Total consumption - 2-step-ahead - Errors relative to AR model and probability of DM test of equal forecasting accuracy (H1: row better than column)

	AR	FCR2	ADL2	TAR2	FCRN2
ME	1.00	-0.14	0.77	0.40	-0.14
MAE	1.00	0.81	0.93	0.82	0.81

	AR	FCR2	ADL2	TAR2	FCRN2
AR		0.93	0.89	0.99	0.93
FCR2	0.07		0.18	0.45	0.48
ADL2	0.11	0.82		0.94	0.81
TAR2	0.01	0.55	0.06		0.54
FCRN	0.07	0.52	0.19	0.46	

Table 5: Errors statistics relative to AR model - Durables - 1-step-ahead - Errors relative to AR model and probability of DM test of equal forecasting accuracy (H1: row better than column)

	AR	FCR	ADL	TAR	FCRN	FCR2	ADL2	TAR2	FCRN2
ME	1.00	0.00	0.44	0.68	0.07	0.05	1.00	0.44	0.06
MAE	1.00	0.91	0.97	1.21	0.91	0.92	1.00	1.07	0.95

	AR	FCR	ADL	TAR	FCRN	FCR2	ADL2	TAR2	FCRN2
AR		0.93	0.68	0.02	0.94	0.90	0.50	0.24	0.81
FCR	0.07		0.19	0.00	0.43	0.28	0.07	0.04	0.05
ADL	0.32	0.81		0.01	0.79	0.74	0.32	0.16	0.59
TAR	0.98	1.00	0.99		1.00	1.00	0.98	0.98	0.99
FCRN	0.06	0.57	0.21	0.00		0.35	0.06	0.03	0.04
FCR2	0.10	0.72	0.26	0.00	0.65		0.10	0.06	0.10
ADL2	0.50	0.93	0.68	0.02	0.94	0.90		0.24	0.81
TAR2	0.76	0.96	0.84	0.02	0.97	0.94	0.76		0.90
FCRN2	0.19	0.95	0.41	0.01	0.96	0.90	0.19	0.10	

Table 6: Errors statistics relative to AR model - Durables - 2-step-ahead - Errors relative to AR model and probability of DM test of equal forecasting accuracy (H1: row better than column)

	AR	FCR2	ADL2	TAR2	FCRN2
ME	1.00	0.07	1.00	0.53	0.06
MAE	1.00	0.91	1.00	1.02	0.91

	AR	FCR2	ADL2	TAR2	FCRN2
AR		0.95	0.50	0.39	0.95
FCR2	0.05		0.05	0.07	0.45
ADL2	0.50	0.95		0.39	0.95
TAR2	0.61	0.93	0.61		0.93
FCRN2	0.05	0.55	0.05	0.07	

Table 7: Errors statistics relative to AR model - Non-durables - 1-step-ahead - Errors relative to AR model and probability of DM test of equal forecasting accuracy (H1: row better than column)

	AR	FCR	ADL	TAR	FCRN	FCR2	ADL2	TAR2	FCRN2
ME	1.00	0.09	1.27	1.24	-0.05	0.08	1.30	1.09	0.10
MAE	1.00	0.82	1.12	1.04	0.99	0.83	1.10	0.96	0.87

	AR	FCR	ADL	TAR	FCRN	FCR2	ADL2	TAR2	FCRN2
AR		0.97	0.08	0.28	0.53	0.96	0.11	0.73	0.88
FCR	0.03		0.01	0.04	0.03	0.33	0.01	0.12	0.23
ADL	0.92	0.99		0.88	0.81	0.99	0.79	0.99	0.97
TAR	0.72	0.96	0.12		0.64	0.95	0.16	0.99	0.92
FCRN	0.47	0.97	0.19	0.36		0.97	0.22	0.59	0.99
FCR2	0.04	0.67	0.01	0.05	0.03		0.02	0.13	0.25
ADL2	0.89	0.99	0.21	0.84	0.78	0.98		0.99	0.96
TAR2	0.27	0.88	0.01	0.01	0.41	0.87	0.01		0.78
FCRN2	0.12	0.77	0.03	0.08	0.01	0.75	0.04	0.22	

Table 8: Errors statistics relative to AR model - Non-durables - 2-step-ahead - Errors relative to AR model and probability of DM test of equal forecasting accuracy (H1: row better than column)

	AR	FCR2	ADL2	TAR2	FCRN2
ME	1.00	0.06	1.05	0.84	0.03
MAE	1.00	0.84	0.99	0.90	0.89

	AR	FCR2	ADL2	TAR2	FCRN2
AR		0.95	0.57	0.97	0.84
FCR2	0.05		0.12	0.27	0.16
ADL2	0.43	0.88		0.90	0.75
TAR2	0.03	0.73	0.10		0.54
FCRN2	0.16	0.84	0.25	0.46	

Table 9: Errors statistics relative to AR model - Semi-durables - 1-step-ahead - Errors relative to AR model and probability of DM test of equal forecasting accuracy (H1: row better than column)

	AR	FCR	ADL	TAR	FCRN	FCR2	ADL2	TAR2	FCRN2
ME	1.00	0.39	0.87	1.21	0.42	0.64	1.06	1.63	0.31
MAE	1.00	1.10	1.07	1.05	1.11	0.99	1.05	1.09	1.15

	AR	FCR	ADL	TAR	FCRN	FCR2	ADL2	TAR2	FCRN2
AR		0.22	0.21	0.19	0.23	0.54	0.22	0.09	0.12
FCR	0.78		0.61	0.67	0.43	0.82	0.67	0.54	0.21
ADL	0.79	0.39		0.60	0.37	0.84	0.62	0.39	0.27
TAR	0.81	0.33	0.40		0.33	0.86	0.50	0.22	0.22
FCRN	0.77	0.57	0.63	0.67		0.82	0.68	0.56	0.32
FCR2	0.46	0.18	0.16	0.14	0.18		0.22	0.09	0.11
ADL2	0.78	0.33	0.38	0.50	0.32	0.78		0.29	0.22
TAR2	0.91	0.46	0.61	0.78	0.44	0.91	0.71		0.33
FCRN2	0.88	0.79	0.73	0.78	0.68	0.89	0.78	0.67	

Table 10: Errors statistics relative to AR model - Semi-durables - 2-step-ahead - Errors relative to AR model and probability of DM test of equal forecasting accuracy (H1: row better than column)

	AR	FCR2	ADL2	TAR2	FCRN2
ME	1.00	0.56	0.87	0.22	0.57
MAE	1.00	1.03	1.04	1.10	1.02

	AR	FCR2	ADL2	TAR2	FCRN2
AR		0.30	0.16	0.11	0.20
FCR2	0.70		0.42	0.27	0.52
ADL2	0.84	0.58		0.23	0.63
TAR2	0.89	0.73	0.77		0.80
FCRN2	0.80	0.48	0.37	0.20	

Table 11: Errors statistics relative to AR model - Services - 1-step-ahead - Errors relative to AR model and probability of DM test of equal forecasting accuracy (H1: row better than column)

	AR	FCR	ADL	TAR	FCRN	FCR2	ADL2	TAR2	FCRN2
ME	1.00	-0.97	0.76	1.06	-0.97	-0.98	0.89	0.51	-0.97
MAE	1.00	1.04	0.96	1.09	1.03	1.02	0.99	1.02	1.00

	AR	FCR	ADL	TAR	FCRN	FCR2	ADL2	TAR2	FCRN2
AR		0.42	0.69	0.10	0.44	0.45	0.53	0.40	0.49
FCR	0.58		0.68	0.37	0.66	0.74	0.60	0.54	0.77
ADL	0.31	0.32		0.04	0.34	0.35	0.11	0.22	0.39
TAR	0.90	0.63	0.96		0.65	0.66	0.94	0.78	0.70
FCRN	0.56	0.34	0.66	0.35		0.59	0.57	0.51	0.66
FCR2	0.55	0.26	0.65	0.34	0.41		0.56	0.50	0.68
ADL2	0.47	0.40	0.89	0.06	0.43	0.44		0.36	0.48
TAR2	0.60	0.46	0.78	0.22	0.49	0.50	0.64		0.55
FCRN2	0.51	0.23	0.61	0.30	0.34	0.32	0.52	0.45	

Table 12: Errors statistics relative to AR model - Services - 2-step-ahead - Errors relative to AR model and probability of DM test of equal forecasting accuracy (H1: row better than column)

	AR	FCR2	ADL2	TAR2	FCRN2
ME	1.00	-0.72	0.70	0.50	-0.72
MAE	1.00	0.90	0.90	0.94	0.86

	AR	FCR2	ADL2	TAR2	FCRN2
AR		0.75	0.92	0.76	0.81
FCR2	0.25		0.49	0.37	0.91
ADL2	0.08	0.51		0.31	0.60
TAR2	0.24	0.63	0.69		0.72
FCRN2	0.19	0.09	0.40	0.28	

7. Conclusions

A relation is often assumed between aggregate consumption and consumer confidence index. Nevertheless, empirical literature results do not confirm this finding so sharply. In some cases it is argued that the predictive ability of confidence would be more important in some phases of the business cycle; moreover, it would be important only when changes in confidence are unusually large. In addition to this, it would seem more reasonable that confidence impact would be more important in cases when good purchases can be more easily shifted in time, as in the case of durables.

All these considerations led me to consider a non-parametric non-linear model to forecast consumption by means of consumer confidence. Moreover, different kind of consumption broken down by durability are taken into account.

Besides the forecasting usefulness, the proposed model reveals itself has a helpful device in order to shed some more light into the seemingly conflicting evidence found in the literature. Indeed, the FCR model estimation results are simple to interpret in this context; for example, in the case of durables an asymmetric threshold effect is found, which leads to predict a sudden fall in this type of consumption as long as the confidence is below a certain threshold, while the converse is not true.

As regards forecasting comparison, a benchmark model is used based only on univariate information coming from different kind of consumption. A linear model and a non-linear non-parametric model are then considered when confidence is included, contemporaneous or lagged one period.

The results show that confidence indicator has a significant predictive ability for durables and non-durables, while it is not useful in the case of semi-durables. Some improvement is visible for services only when lagged confidence is considered. When forecasting directly the total ,consumption it turns out that confidence remains useful.

Differences emerge when considering the linear ADL model and the non-linear non-parametric model For the total and durables, the latter reaches a better performance than the linear one, although not significantly at conventional levels. In the case of non-durables only the FCR model beats the benchmark, while the ADL has significantly worse performance.

Overall, the results confirm that confidence indicator has a predictive content, even though its relevance is probably overvalued. In any case, the use of a non-linear model can help in picking-up better its predictive content for consumption, especially when some disaggregate series are the object of the analysis. Moreover, the FCR model reveals itself also as a nice descriptive device, which can overcome the rigidity of linear models and avoid the difficulties related in the choice of a class of non-linear parametric model.

References

Acemoglu, D. and Scott, A. (1994). Consumer confidence and rational expectations: Are agents' beliefs consistent with the theory? *The Economic Journal*, 104(422):1–19.

Al-Eyd, A., Barrell, R., and Davis, E. P. (2009). Consumer confidence indices and short-term forecasting of consumption. *Manchester School*, 77(1):96–111.

Batchelor, R. and Dua, P. (1998). Improving macro-economic forecasts: The role of consumer confidence. *International Journal of Forecasting*, 14(1):71–81.

Bram, J. and Ludvigson, S. (1998). Does consumer confidence forecast household expenditure? a sentiment index horse race. *Economic Policy Review*, pages 59–78.

Cai, Z., Fan, J., and Yao, Q. (2000). Functional-coefficient regression models for nonlinear time series. *Journal of the American Statistical Association*, 95(451):941–956.

Carroll, C. D., Fuhrer, J. C., and Wilcox, D. W. (1994). Does consumer sentiment forecast household spending? If so, why? *The American Economic Review*, 84(5):1397–1408.

Chow, G. C. and Lin, A.-l. (1971). Best linear unbiased interpolation, distribution, and extrapolation of time series by related series. *The Review of Economics and Statistics*, 53(4):372–375.

Desroches, B. and Gosselin, M.-A. (2002). The usefulness of consumer confidence indexes in the United States. Working Papers 02-22, Bank of Canada.

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

Fan, J., Yao, Q., and Cai, Z. (2003). Adaptive varying-coefficient linear models. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 65(1):57–80.

Fuhrer, J. C. (1988). On the information content of consumer survey expectations. *The Review of Economics and Statistics*, 70(1):140–144.

Garner, C. A. (2002). Consumer confidence after September 11. *Economic Review*, (Q II).

Harvill, J. L. and Ray, B. K. (2005). A note on multi-step forecasting with functional coefficient autoregressive models. *International Journal of Forecasting*, 21:717–727.

Katona, G. (1971). Consumer durable spending. *Brookings Papers on Economic Activity*, 1971(1):234–239.

Ludvigson, S. C. (2004). Consumer confidence and consumer spending. *Journal of Economic Perspectives*, 18(2):29–50.

Malgarini, M. and Margani, P. (2007). Psychology, consumer sentiment and household expenditures. *Applied Economics*, 39(13):1719–1729.

Matsusaka, J. G. and Sbordone, A. M. (1995). Consumer confidence and economic fluctuations. *Economic Inquiry*, 33(2):296–318.

McIntyre, K. H. (2007). Reconciling consumer confidence and permanent income consumption. *Eastern Economic Journal*, 33(2):257–275.

Proietti, T. and Frale, C. (2007). New proposals for the quantification of qualitative survey data. CEIS Research Paper 98, Tor Vergata University, CEIS.

Qiao, Z., McAleer, M., and Wong, W.-K. (2009). Linear and nonlinear causality between changes in consumption and consumer attitudes. *Economics Letters*, 102(3):161–164.

Roos, M. W. (2008). Willingness to consume and ability to consume. *Journal of Economic Behavior & Organization*, 66:387–402.

Teräsvirta, T., van Dijk, D., and Medeiros, M. C. (2005). Linear models, smooth transition autoregressions, and neural networks for forecasting macroeconomic time series: A re-examination. *International Journal of Forecasting*, 21(4):755–774.

Throop, A. W. (1992). Consumer sentiment: its causes and effects. *Economic Review*, pages 35–59.

Xia, Y. and Li, W. K. (1999a). On single-index coefficient regression models. *Journal of the American Statistical Association*, 94(448):1275–1285.

Xia, Y. and Li, W. K. (1999b). On the estimation and testing of functional-coefficient linear models. *Statistica Sinica*, 9:735–757.

A. List of questions composing the consumer confidence

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

- got a lot better
- got a little better
- stayed the same
- got a little worse
- got a lot worse
- don't know.

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

- get a lot better
- get a little better
- stay the same
- get a little worse
- get a lot worse
- don't know.

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

- got a lot better
- got a little better
- stayed the same
- got a little worse
- got a lot worse
- don't know.

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

- get a lot better
- get a little better
- stay the same
- get a little worse
- get a lot worse
- don't know.

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

- increase sharply
- increase slightly
- remain the same
- fall slightly
- fall sharply
- don't know.

Compared to the past 12 months, do you expect to spend more or less money on major purchases (furniture, electrical/electronic devices, etc.) over the next 12 months? I will spend...

- much more
- a little more
- about the same
- a little less
- much less
- don't know.

In view of the general economic situation, do you think that now is...?

- a very good moment to save
- a fairly good moment to save
- not a good moment to save
- a very bad moment to save
- don't know.

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

- very likely
- fairly likely
- not likely
- not at all likely
- don't know.

Which of these statements best describes the current financial situation of your household?

- we are saving a lot
- we are saving a little
- we are just managing to make ends meet on our income
- we are having to draw on our savings
- we are running into debt
- don't know.