27th CIRET Conference, Warsaw, September 2004

Session:
6B
Use of Survey Data or Cyclical Indicators for Micro- or Macroeconomic Analysis

Consumer Surveys and Reality

Maurizio Bovi

ISAE*

Abstract

This paper investigates the usefulness of Italian consumer surveys as estimation and forecasting tool over the period 1982-2003. To this end, standard consumption equations are estimated and then compared, in terms of in-sample and out-of-sample predictive ability, with corresponding models which differ from them only because of the presence of the confidence indicator. Unlike mainstream literature, the present work focuses on the relationships between subjective and objective information at a less aggregate level. In particular, the overall sentiment index is divided into four sub-indices related to the opinion about the i) current, ii) future, iii) general, and iv) personal situation. In turn, the total private consumption is divided in five items. The idea behind is to check if one attitudinal measure is more or less informative than another, and if some outlay is more or less "sentiment sensitive" than another. It is shown that the qualitative information obtained from household surveys improves both the goodness-of-fit of consumption equations and their forecasting performances. It is noteworthy that these improvements are all the more evident when working on disaggregated data, *i.e*., linking a particular kind of consumption to a particular sub index. For instance, perceptions about the future help to explain consumption for services more than disbursement for non durable goods.

JEL classification : C22, C52, D12, E21, E27.

Keywords: consumer sentiment; surveys; Consumption; time series.

* Affiliation: Institute for Studies and Economic Analyses (ISAE), Address: Piazza dell'Indipendenza, 4 – 00185 Roma – Italy; e-mail: m.bovi@isae.it., Fax #: +39.06.4448.2249, Phone #: +39.06.4448.2751.

1. Introduction

Surveys of people in households add a psychological dimension that can be tentatively quantified for use in the study of economic systems. In fact, the effect of consumers attitudes on economic activity is a subject of great and growing interest among policymakers and economists. For instance, the worsening of households confidence has often been reported as one of the main causes of the early '90s world-wide recession. The peculiar nature of this subjective information suggests to accurately analyse its advantages and drawbacks, on the one hand, and a careful use of it, on the other hand. Theoretically, the psychological dimension has never had a relevant role in economic literature. Excluding the attempts by Katona (1951) and the Keynesian allusion to "animal spirits", households' and firms' behaviours have always been studied without considering their emotional attitude to act. Needless to say, it is very hard to gather reliable information about this latter. The mainstream solution is based on directly interviewing agents. The results of these surveys are then arranged into "balances", which are the differences between the percentages of positive and negative answers. The Italian consumers' sentiment index monthly worked out by ISAE[1] is based on a representative sample of 2,000 respondents, and concerns the consumers' expectations on the overall economic situation and on their own financial situations. It is computed as the arithmetic average of the (seasonally adjusted) net balances of positive minus negative responses to nine questions on: (1) anticipated personal financial conditions over the coming year; (2) anticipated economic conditions over the coming year; (3) anticipated job availability conditions; (4) whether now is a good or bad time to save; (5) personal financial conditions over the past year; (6) economic conditions over the past year; (7) personal opinion regarding the household's budget; (8) personal saving possibilities over the coming year; (9) whether now is a good or bad time to buy major household items.

Provided that consumers' sentiment indices (CS) have additional information content about the economic situation beyond that already contained in quantitative variables, they can be very helpful because of the timeliness of their availability. The ISAE index is diffused in the last ten days of the reference month, while the promptest quantitative measure (the GDP preliminary estimate) is released by the Italian National Institute of Statistics (ISTAT) with a delay of forty days. Despite the popularity of the qualitative indices, there is little consensus about their ability to collect information on consumer spending that is not already captured by economic fundamentals. In response to the widespread belief that consumers' opinions and expectations influence the direction of the economy, a growing number of studies have set out to analyze the relationship between consumer attitudes and economic variables. So far, the literature on the predictive power of CS have showed mixed conclusions (Leeper, 1992; Carroll, Fuhrer, and Wilcox, 1994; Matsusaka and Sbordone 1995; Berg and Bergstrom, 1996; Batchelor and Dua, 1998; Bram and Ludvigson, 1998; Golinelli and Parigi, 2003).

In order to reflect and represent total economic activity, surveys data are used to calculate a composite indicator. Thus in literature it is often compared with the performance of a reference variable which is also all inclusive. The GDP is the obvious choice, but the CS has often been related to total private consumption as well. However, it can be useful to invert the process, that is to collate subjective information with objective one at a less aggregate level. In that is the main novelty of this paper, which is the first formal empirical investigation of household attitudes that associates four different sub-indices with six categories of consumer spending. The idea behind is to check if one emotional measure is more/less informative than another, and if some outlay is more/less "sentiment sensitive" than another. In the former case it may happen that even if the overall index turns out to be uninformative, it could be not true for some of its components. In the latter case one can expect that a serene mood could lead to an increase in the consumption for, *e.g.*, recreation, while expenses incurred for medical care should be driven by psychological motivations only, if any, to a lesser extent.

From the methodological point of view I use a two-step procedure to determine the in-sample and out-of-sample informative power of CS. First, I consider a standard baseline equation for consumption growth that does not include attitudinal survey measures. Then I add consumer sentiment indicators to

[1] For a description of the ISAE Consumers' confidence indicator, see Martelli (1998).

the baseline equation and test which CS, if any, improve the goodness-of-fit (adjusted R-squared) and/or the root mean square error (RMSE) of the benchmark model. As in most of the analyses in the field (for an exemption, see Golinelli and Parigi, 2003) the statistical framework is univariate. This may raise simultaneity bias issues, but in the present empirical design the Johansen procedure suggests the existence of, at most, only one cointegration vector. This means that the single-equation approach can be validly pursued because the weak exogeneity requisite holds (Engle *et al.*, 1983). Within the single equation framework, I follow the general-to-specific approach in its Wickens-Breusch (1988) form which i) increases the efficiency of the estimates; ii) allows testing the statistical significativity of the error correction term in a single step; iii) is particularly suited when dealing with both theoretical and empirical information. Also, since the residuals are homoschedastic innovations and the parameters are constant, the baseline models are empirically admissible and coherent (Hendry, 1995). Therefore they are both reliable and strong contenders for the CS-augmented models.

Results show that the qualitative information obtained from household surveys improves both the goodness-of-fit of consumption equations and their forecasting performances. These improvements are all the more evident when working on disaggregated data, *i.e*. linking a particular kind of consumption to a particular sub index. For instance, consumption for services seems to be sensitive to households' personal conditions and expectations in the sense that (alternatively) adding these sub-indices in the relative baseline model allows improving both the in sample and the out of sample performances of the baseline equation. This result does not hold for other indices, in particular for the most aggregate one which, actually, shows the lowest predictive ability. As refer to the health care outlay, psychological measures result useless in forecasting exercises. It can be seen as an indirect clue that in order to better understand the informative content of CS a careful treatment of macroeconomic data is needed. Furthermore, as already pointed out by literature, subjective indicators help to shed some light about extraordinary economic episodes. Sub-sample analyses show that throughout the two-recessions period 1992-2003 some CS sub-indices increase their forecasting ability while, once again, it does not hold for the overall indicator. Finally, the out-of-sample experiments performed over the last four years (2000-2003) show better results when using some sub indices. Thus, even during the recent puzzling behaviour of the Italian climate measures, languishing in a very low level as opposed to a relatively sustained growth in the total consumption, a disaggregated analysis is still able to shed some lights about the predictive ability of qualitative measures.

The paper is organised as follow. The next section describes the surveys data, Sections 3 and 4 outline the empirical framework, Section 5 collects the findings. Concluding remarks close the paper.

2. Surveys Data

The qualitative data are drawn by ISAE and cover the sample period 1982 to 2003. Since macroeconomic data are quarterly, I converted the monthly frequency of ISAE data by a simple average. As mentioned, the overall index (1980=1) is formed by nine questions. Following the ISAE procedure, I exhaustively aggregate these latter into two pairs of sub-indices. The first couple deals with present and future situations; the other pair distinguishes between personal and economic conditions in the country as a whole. More in detail, the questions are collected in the following way:

- Personal:

1) anticipated personal financial conditions over the coming year;
4) whether now is a good or bad time to save;
5) personal financial conditions over the past year;
7) personal opinion regarding the household's budget;
8) personal saving possibilities over the coming year;
9) whether now is a good or bad time to buy major household items;

- General

2) anticipated economic conditions over the coming year;
3) anticipated job availability conditions;
6) economic conditions over the past year.

- Present:

4) whether now is a good or bad time to save;
7) personal opinion regarding the household's budget;
9) whether now is a good or bad time to buy major household items;
5) personal financial conditions over the past year;
6) economic conditions over the past year;

- Future:

1) anticipated personal financial conditions over the coming year;
2) anticipated economic conditions over the coming year;
3) anticipated job availability conditions;
8) personal saving possibilities over the coming year.

Clearly, the indices Present and Personal on the one side, and General and Future on the other side, are highly correlated because they share some items. The following figures 1-4 report the (seasonally adjusted) time series of the five indices (the overall and the four sub-indices).

Figure 1. Italian Consumer Sentiment Indices (1980=1, seasonally adjusted). Overall vs Personal

Figure 2. Italian Consumer Sentiment Indices (1980=1, seasonally adjusted). Overall vs General

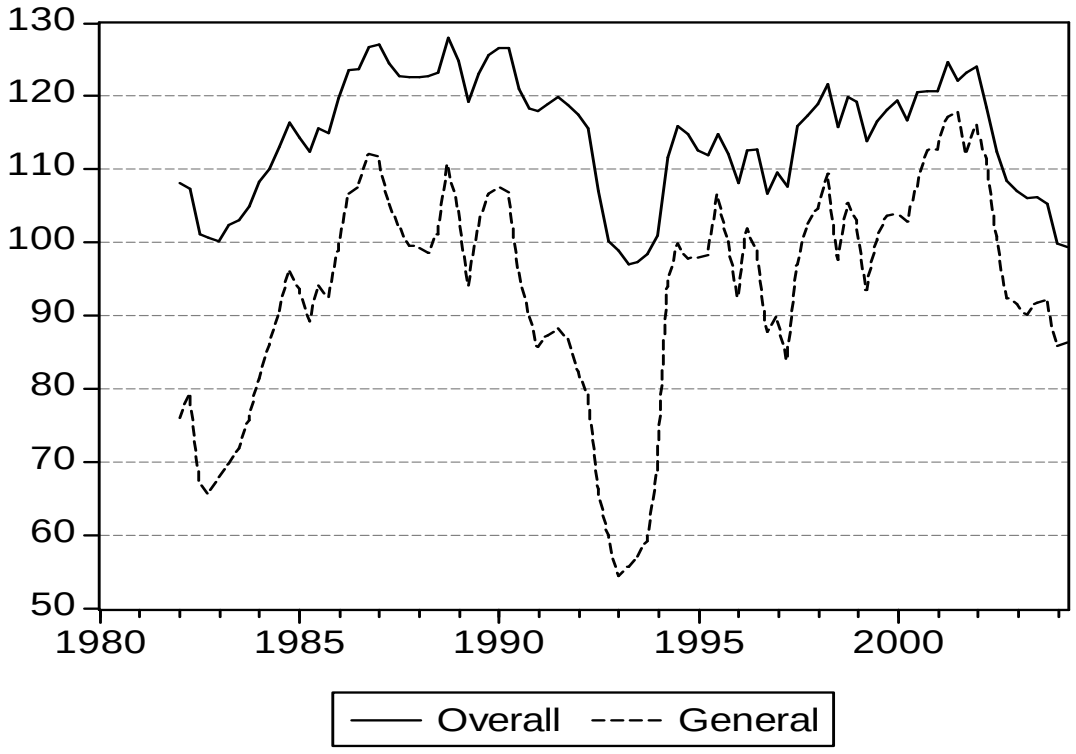


Figure 3. Italian Consumer Sentiment Indices (1980=1, seasonally adjusted). Overall vs Present

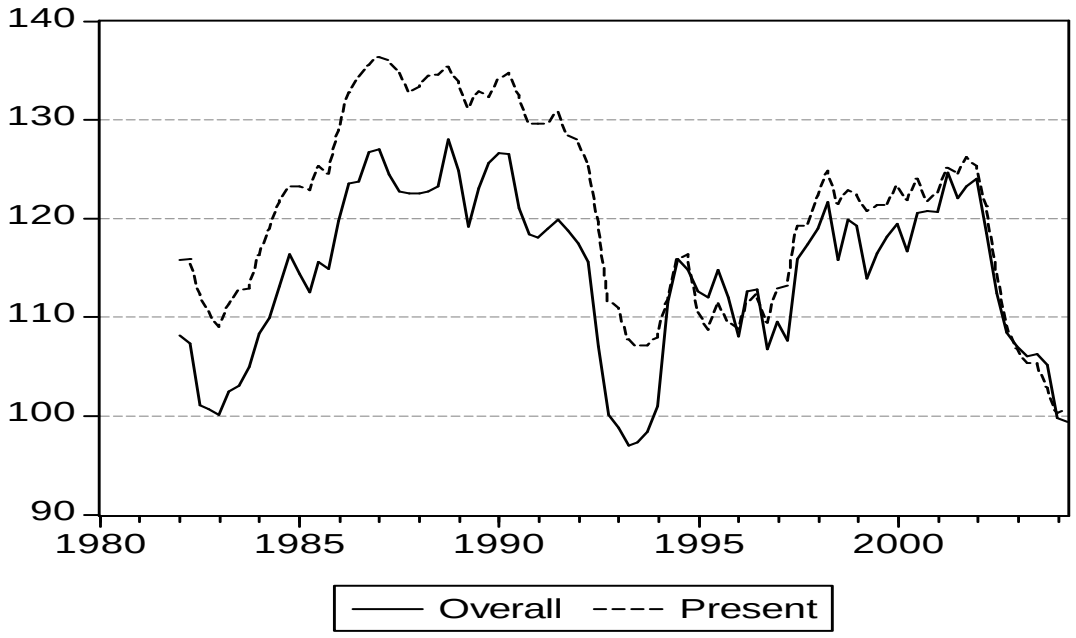
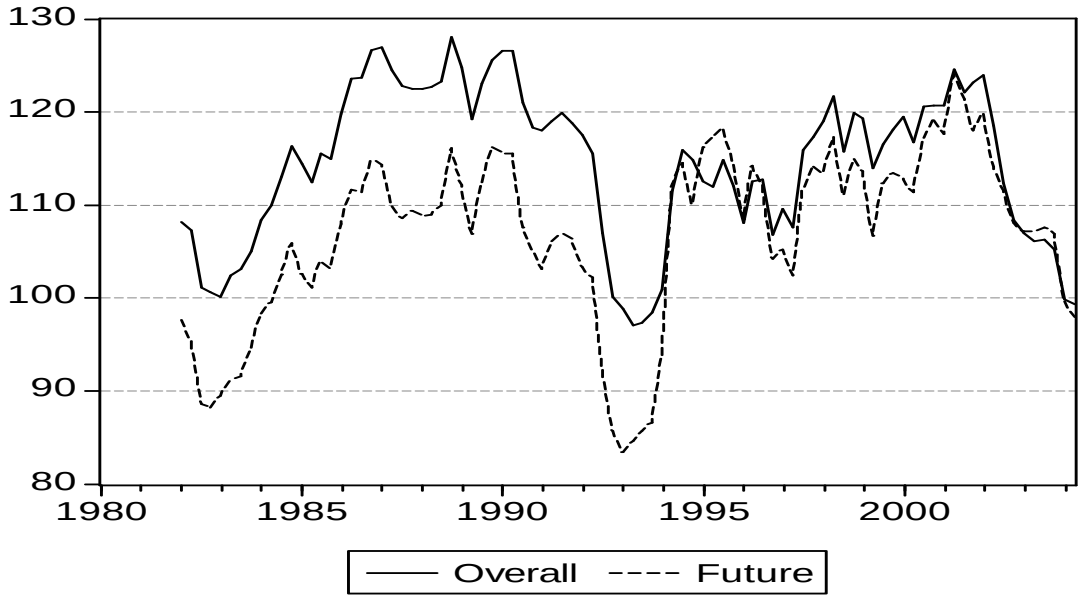


Figure 4. Italian Consumer Sentiment Indices (1980=1, seasonally adjusted). Overall vs Future



A visual inspection of the figures 1-4 highlights some stylised fact. First, the indices Personal and Present have always been higher than the Overall one. The opposite holds for the other two complementary measures. Second, starting from the mid-90s there seems to be a tendency to the reduction of the gap between indices. This could be partly due to the change in the methodological approach to the households - now interviews are made by phone, while they were direct before the 1995. Third, and more interestingly, General and Future show wider reactions to the early 90s recession and then a positive trend up to the present slowdown in the economic activity. On the contrary, Personal and Present remained more constant in the last decade, while show a stronger reduction in the recent years. As a consequence, these indices languish now in an historical minimum, whereas the lowest values for General and Future were reached in 1993. A tentative explanation may be found in the European commitments signed at Maastricht in 1992. In fact, in that year there was the last devaluation of the Lira, while henceforth there was a worsening of the fiscal stance. May be that people felt the devaluation as a deterioration of the system more than of their personal condition (in the overall index there is not a question referring to the inflation), while the protracted fiscal consolidation (and the following impossibility to implement counter-cyclical policies) impinged especially on the individual position. For instance, the strong reduction in the interest rates caused by the fiscal stance have had a dramatic impact on the disposable income of Italian families. Also, the public sector reduced its role as employer of last resort. Altogether this suggests that each sub index can tell a different, and potentially useful, story about the relationships between macroeconomic variables and people's feelings. This calls for the deeper and more formal analyses which I perform in the next sections.

3. Macroeconomic Data and Preliminary Econometric Analysis

Economic literature pointed out several determinants of consumption. The path-breaking paper by Davidson *et al*'s (1978) specifies two causes of long-run consumption: income and inflation. However, alternative consumption theories and empirical evidence from advanced industrial countries suggest that consumption is likely to be influenced by additional variables, such as wealth, demographic factors, liquidity constraints, and uncertainty (Muellbauer and Lattimore, 1995). The timing of the variables has been investigated as well. Nelson (1987) showed that consumption growth is correlated with lagged growth in disposable income. Working on the famous Hall's (1978) random walk hypothesis, Campbell and Mankiw (1989, 1990) have argued that consumption growth is a random walk once the response of consumption growth to the contemporaneous change in income is taken into account.

In an attempt to mediate between data availability, degrees of freedom and theoretical suggestions, I select a set of "standard" regressors, namely: the real wage; the real interest rate (loans to households); the annual rate of inflation. I then add the (seasonally adjusted) number of new registration of cars (Car) because it is used by ISTAT to estimate durable goods quarterly data. These determinants are used to model several kinds of consumption[2]. A widespread used way to disentangle the total consumption of households deals with the durability of the goods, which gives consumption for durable goods, non durable goods and services. Then, according to the goal of this paper I analyse two other kinds of consumption. The first wishes to insulate the consumption for "happiness" and to this end I sum items such as recreation, hotels, and "abroad" (a proxy for holidays). The second tries to define tentatively the expenses for "illness" and it is formed by disbursements for health care. The logic behind is that a serene mood could lead to an increase in the consumption for recreation, while outlays incurred for medical care should be driven by psychological motivations only, if any, to a lesser extent. In other words the idea is to check if the information content of subjective indicators depends on the type of the macroeconomic variables under analysis. To sum up, I have six baseline models according to six kinds of consumption (total, durable goods, non durable goods, services, happiness, and illness), and I use four right-hand-side variables (**Z**= real wages, real interest rate, inflation rate, new registration of cars).

Given the aim of the work and the use of a very general unrestricted Autoregressive Distributed Lag (ADL) model as a starting point, my approach is "one-model-fits-all". For instance, it is obvious that the expenditure for Illness should depend on demographic factors as well. However for the comparisons of section 5, it is sufficient that the models result an empirically admissible and coherent representation of the data.

Before detailing the baseline models (see section 4), it is important to recall that the first necessary step to validly estimate a model is the univariate analysis of the stochastic properties of the series involved. The attention devoted to this topic is well deserved for several reasons. First, in contrast to stationary or trend stationary time series, models with a stochastic trend have time dependent variances that go to infinity with time, thus they are persistent in the sense that shocks have permanent effects on the values of the process. Second, when a series is used in regressions with other variables the interpretation of the regression results can depend on whether the variables involved are trend (TS) or difference stationary (DS). This phenomenon is related to the "nonsense" and "spurious" regression literature due to Yule (1926) and Granger and Newbold (1974). It is also well known that unit root tests are based on asymptotic critical values. One expects in finite samples that the use of asymptotic critical values will result in over-rejection. I address this problem by performing two unit root tests. The first (NP) was worked out by Ng and Perron (2001). It yields both substantial power gains and a lower size distortions over the standard

[2] As will be explained in the text, I use one only general unrestricted model for each kind of consumption, thus the variable "Car" enters all the consumption models (but Illness). The logic of the inclusion is that some non durable and services expenditures are related to the ownership and the use of a car (insurance, gasoline, maintenance, etc.).

unit root tests, maintaining the null of unit root. NP offer four test statistics based on the GLS detrended data y_t^d. Altogether these statistics are modified versions of Phillips-Perron $Z_\square$ and Z_t statistics (1988), the Bargava (1986) R_1 statistic, and the Elliot *et al.* Point Optimal statistic (1996):

$$MZ_\alpha = (T^{-1}(y_t^d)^2 - f_0)/2\kappa \quad [1]$$

$$MSB = (\kappa/f_0)^{1/2} \quad [2]$$

$$MZ_t = MZ_\alpha \text{ X } MSB \quad [3]$$

$$MPT = \bar{c}^2 \kappa - \bar{c}T^{-1}(y_t^d)^2)/f_0 \quad \text{(if exogenous = constant)} \quad [4]$$

$$MPT = \bar{c}^2 \kappa + (1-\bar{c})T^{-1}(y_t^d)^2)/f_0 \quad \text{(if exogenous = constant, trend)} \quad [5]$$

where $\kappa = \sum_{t=2}^{T}(y_{t-1}^d)^2)/T^2$ and f_0 is an estimate of the residual spectral density at the zero frequency[3]. The choice of the autoregressive truncation lag, p, is critical for correct calculation of f_0. Here p is chosen by using the modified AIC suggested by Ng and Perron (2001).

The second is the KPSS test (Kwiatkowski *et al.* (1992)), which can be thought as complementing the NP one because it tests the null hypothesis that the variable follows a TS stochastic process. Suppose the NP test fails to reject the unit root null because of low power. The KPSS test which has (trend) stationarity as the null should indicate the data have no unit roots. On the other hand, if the KPSS test rejects the trend stationarity null, then we have stronger evidence for unit root persistence. That is, consistent results from NP and KPSS tests yield more persuasive evidence on data persistence, while conflicting results indicate uncertainty associated with the interpretation of the individual test outcomes. The KPSS test is based upon the residuals from the OLS regression of y_t on the exogenous variables x_t:

$$y_t = x_t'\delta + u_t \quad [6]$$

The LM statistic is be defined as:

$$LM = \sum_t^n S(t)^2/(T^2 f_0) \quad [7]$$

where f_0 is an estimator[4] of the residual spectrum at frequency zero and where S(t) is a cumulative residual function:

$$S(t) = \sum_{r=1}^{t} \hat{u}_r \quad [8]$$

based on the residuals $\hat{u} = yt - x'\hat{\delta}(0)$. I maintain the same lag length selection criterion already used in the NP test.

[3] The frequency zero spectrum method used is the AR-GLS detrended.

[4] The frequency zero spectrum method used is the Kernel-Bartlett sum-of-covariances.

The following table 1 collects the empirical results.

Table 1. Unit root tests (quarterly data 1980.1-2003.4)

	MZ_α	MZ_t	MSB	MPT	KPSS
Total consumption	-6.75826	-1.70761	**0.25267**	**13.6021**	0.179203
Δ(Total consumption)	**-21.1496**	**-3.24773**	0.15356	4.33431	**0.064255**
Durables	-8.05009	-1.97502	**0.24534**	**11.4133**	0.118274
Δ(Durables)	-3.62172	-1.34559	**0.37153**	**25.1592**	**0.064629**
Non Durables	-5.98427	-1.63820	**0.27375**	**15.1369**	0.174084
Δ(Non Durables)	-2.71788	-1.02638	**0.37764**	**29.2161**	**0.109472**
Services	-2.79698	-0.93964	**0.33595**	**25.8587**	0.268486
Δ(Services)	**-17.9541**	**-2.98072**	**0.16602**	**5.17083**	**0.050400**
Happiness	1.07461	0.98842	**0.91980**	**60.9777**	1.270553
Δ(Happiness)	-0.58889	-0.40363	**0.68541**	**26.1136**	**0.086438**
Illness	-3.80728	-1.28162	**0.33662**	**22.6197**	0.192348
Δ(Illness)	-0.46344	-0.47952	**1.03469**	**195.224**	**0.136969**
Real wage	0.76764	0.56129	**0.73119**	**38.8100**	1.045664
Δ(Real wage)	-3.55903	-1.08201	**0.30402**	**6.89841**	**0.103472**
Real interest rate	-2.19994	-0.88833	**0.40380**	**33.7675**	0.280794
Δ(Real interest rate)	**-42.4318**	**-4.60587**	0.10855	2.14864	**0.027553**
Inflation	-13.5926	-2.59510	**0.19092**	**6.77348**	1.116020
Δ(Inflation)	**-36.1758**	**-4.24189**	0.11726	2.58059	**0.074645**
Car	-9.09355	-2.12656	**0.23385**	**10.0447**	0.329684
Δ(Car)	-6.17382	-1.75653	**0.28451**	**14.7597**	**0.059292**
Overall sentiment	-5.32131	-1.45439	**0.27331**	**16.5535**	**0.098587**
Personal	-4.47922	-1.24193	**0.27726**	**18.4745**	**0.141919**
General	-6.88406	-1.80451	**0.26213**	**13.2966**	**0.084934**
Future	-8.91567	-1.98958	**0.22316**	**10.6821**	**0.070759**
Present	-6.01753	-1.58465	**0.26334**	**15.0102**	**0.105081**

Lag length criterion: Modified AIC; constant and trend included. All variables, but the interest rate, are in log. Car = # of new registration of cars. Happiness = Consumption for recreation, hotels, restaurants and "abroad". Illness = Consumption for health care. Bold values imply stationariety (5%), that is imply the rejection of the null of unit root or the non rejection of the null of stationariety (in the KPSS case). $\Delta X_t \equiv X_t - X_{t-1}$.

As usual, when dealing with short samples results about the statistical nature of the time series are mixed. However, the overall picture emerging from Table 1 suggests that the macroeconomic variables seems to be I(1), while indices seems to be stationary. This latter finding is somewhat expected because it is hard to think about an everlasting "irrational exuberance (or apprehension)" in all agents. Since the quantitative (economic) time series are integrated of the same order, the next step is to test whether them (alternatively each type of consumption and its above mentioned quantitative[5] determinants) are cointegrated. In order to save space I do not report the results of Johansen procedure which seems to suggest that, at most, there is only one cointegration vector. This means that the single-equation approach can be validly pursued because the weak exogeneity requisite holds (Engle *et al*., 1983).

[5] Given that qualitative indices result stationary they can not enter the ECM (see Section 4).

4. The baseline models

Within the single-equation framework, I follow the general-to-specific approach in its error-correction (ECM) version to modelling consumption (Davidson *et al*., 1978). The major feature of the ECM is that it distinguishes between short term and long term effects. The specification of long-run components in the consumption function draws upon economic theory, while dynamics are a clear case of "measurement without theory", at which Hendry (1995) refers to the "let the data speak" approach. This is another reason, beyond their stochastic properties, why I do not enter the CS indices in the error correction term - they are measurement without theory. This statistical background seems to be quite suitable in the present context with mixed information coming from both empirical and theoretical applications. The starting "general" system is an ADL(4,4) reparametrised *a là* Wickens-Breusch (1988) in order to estimate (in one single step and more powerfully in small samples) the possible presence of the ECM. Formally (i=0,…,4; j=1,…,4):

$$\Delta C_t = \alpha + \Sigma\gamma\Delta\mathbf{Z_{t-i}} + \Sigma\eta\Delta C_{t-j} + \beta\mathbf{Z_{t-1}} + \delta C_{t-1} + \varepsilon_t \qquad [9]$$

where $\Delta X_t \equiv X_t - X_{t-1}$, C_t is the consumption variable, and $\mathbf{Z_t}$ is the right-hand-side variables vector as defined in section 3. To account for monetary illusion in the short run, I enter the first differences of the interest rate and of the (log)wage in nominal term. Diagnostic tests suggest that four lags are sufficient for a reliable general unrestricted model to start with. After having obtained the results for the whole set of the variables (the "general" model), I sequentially delete the insignificant variables to select a parsimonious (specific) model for increasing the precision of the estimates (Hendry, 1995). The specification search ends according to two information criteria[6] based on the log-likelihood value. The first one, proposed by Akaike (1973) penalizes the log-likelihood by 2n/T for n parameters and a sample size of T; the second (Schwarz, 1978), by n log(T)/T.

As requested in order to see if the model captures the essential characteristic of the data (denoted congruence), in each step of the testing down procedure I check if i) the parameters are constant (via a recursive analysis); ii) the residuals are homoschedastic innovations. All in all it results that the baseline models are empirically admissible and coherent. Thus, in the next section I will compare CS-augmented models with very strong and reliable contenders. Some diagnostic tests referring to all these models are reported in the following section (see table 2).

[6] Variables not significant at the conventional levels (5%) are considered irrelevant and excluded from the regression. At the end of the specification search I performed an F-test on the joint significativity of the deleted variables. Evidence suggests to proceed with the "specific" model. All empirical results are available upon request from author.

5. The Information Content of Consumer Sentiment Indices

To investigate the usefulness of Italian consumer surveys as estimation and forecasting tool over the period 1982.1-2003.4, I estimate new models which differs from the baseline only because of the inclusion of alternative contemporaneous and lagged confidence indicators. According to the results of section 3, CS indices are stationary processes, thus they enter the baseline models in log-levels. As for the other determinants, four lags proved to be sufficient for a reliable general unrestricted model to start with. Then I repeat the testing down procedure. Finally, I compare each benchmark model with the corresponding sentiment-augmented in their specific form in terms of goodness-of-fit (adjusted R^2) and out-of-sample forecast performance statistics (root-mean-square % error). These latter are computed by estimating the specific model over the period 1980.1-1999.4, and then by performing both static and dynamic forecasts over the next sixteen quarters (up to 2003.4). Table 2 collects the findings of all experiments.

Table 2. Standard vs Consumers Sentiment Augmented Consumption Growth Models

		Models of Total Consumption							
Sample		*Baseline*	+ Overall	+ Present	+ Future	+ General	+ Personal	+ Pres+Fut.	+ Gen+Pers
1980-03		*0.68628*	0.68185	**0.68932**	0.68299	------	**0.71101**	0.68448	**0.70838**
1987-03	**Adj-R^2**	*0.69367*	0.68006	**0.69604**	0.68324	------	**0.69911**	0.68303	**0.69438**
1992-03		*0.70079*	0.68000	**0.71558**	0.68372	------	**0.74194**	**0.70500**	**0.76010**
2000-03	**RMSPE_s**	*0.41075*	0.43506	0.42196	0.45071	------	0.41118	0.44582	0.43037
2000-03	**RMSPE_d**	*0.86358*	1.24603	**0.85639**	1.17518	------	0.88937	1.00572	0.88808
		Models of Consumption for Durable Goods							
80.1-03.4		*0.86882*	------	------	**0.87429**	**0.87225**	**0.87303**	**0.88208**	**0.88157**
87.1-03.4	**Adj-R^2**	*0.87101*	------	------	**0.87881**	**0.87241**	**0.87841**	**0.88511**	**0.88002**
92.1-03.4		*0.89576*	------	------	**0.91646**	0.89416	**0.90138**	**0.90070**	0.89348
00.1-03.4	**RMSPE_s**	*1.17977*	------	------	1.30916	1.23579	1.40113	**1.17287**	**1.14684**
00.1-03.4	**RMSPE_d**	*2.11762*	------	------	2.72201	**1.77434**	2.47737	**1.97769**	**2.06919**
		Models of Consumption for Non-Durable Goods							
80.1-03.4		*0.45967*	**0.45993**	**0.46477**	**0.46388**	------	**0.48301**	**0.48087**	**0.51496**
87.1-03.4	**Adj-R^2**	*0.44770*	0.43897	0.41167	0.42465	------	0.39069	0.42286	**0.54438**
92.1-03.4		*0.38288*	**0.38853**	0.38195	0.35771	------	0.37636	0.34416	0.23787
00.1-03.4	**RMSPE_s**	*0.61607*	0.68117	0.80861	0.72136	------	0.64605	**0.54664**	0.64906
00.1-03.4	**RMSPE_d**	*1.01191*	1.60196	2.09897	1.64914	------	1.13034	**0.77583**	**0.89510**
		Models of Consumption for Services							
80.1-03.4		*0.38368*	0.37394	**0.39074**	**0.38955**	0.37713	**0.39543**	**0.40650**	**0.39972**
87.1-03.4	**Adj-R^2**	*0.33333*	0.30451	0.33329	**0.35948**	**0.34316**	**0.35000**	**0.35092**	**0.34690**
92.1-03.4		*0.43385*	**0.47593**	**0.52668**	**0.43832**	**0.47651**	**0.46776**	0.38137	**0.43655**
00.1-03.4	**RMSPE_s**	*0.35411*	0.42201	0.50067	0.45292	0.53487	0.37995	0.51405	0.41917
00.1-03.4	**RMSPE_d**	*1.09121*	**1.01232**	1.20677	**0.94159**	1.23044	**0.80676**	1.33975	**0.90666**
		Models of Consumption for Happiness							
80.1-03.4		*0.36959*	------	0.36677	**0.38009**	0.35287	**0.38307**	**0.41198**	**0.38851**
87.1-03.4	**Adj-R^2**	*0.25953*	------	**0.27455**	**0.28592**	**0.28745**	**0.29088**	**0.40523**	**0.32817**
92.1-03.4		*0.29642*	------	**0.32362**	**0.40172**	**0.36660**	**0.32998**	**0.31313**	**0.44910**
00.1-03.4	**RMSPE_s**	*1.20972*	------	**1.11277**	**1.19431**	1.39082	**1.18250**	1.31710	1.26380
00.1-03.4	**RMSPE_d**	*2.07997*	------	2.33869	2.15515	2.31238	**2.01811**	2.31472	2.36184
		Models of Consumption for Illness							
80.1-03.4		*0.352144*	------	------	------	------	------	------	------
87.1-03.4	**Adj-R^2**	*0.362282*	------	------	------	------	------	------	------
92.1-03.4		*0.387924*	------	------	------	------	------	------	------
00.1-03.4	**RMSPE_s**	*2.086742*	------	------	------	------	------	------	------
00.1-03.4	**RMSPE_d**	*9.229285*	------	------	------	------	------	------	------

Happiness = Consumption for recreation, hotels, restaurants and abroad. Illness = Consumption for health care.
All variables, but the interest rate, are logged. The reported statistics are drawn from the specific models obtained after the testing down procedure. In all cases, the general model is the ADL(4,4) reported in the text as equation [9].
The general baseline model includes: real wage, real interest rate, inflation rate, new registration of cars. The general CS-augmented models just add to it the contemporaneous and the first four lags of the log-level of the index/indices reported in the second row ("+ Overall"; "+ Present" etc.). RMSPE = root-mean-square % error. Tag=_s refers to static forecasting exercises; tag=_d refers to dynamic forecasting exercises. They are computed on sixteen-steps (2000.1-2003.4) out-of-sample forecasts (of the log-level of the dependent variable) based on regressions over the period 1980.1-1999.4. Bold values emphasise lower R2 and higher RMSPE realised by the baseline model which, in turn, implies that climate indices have information content other than that contained in macroeconomic variables. The other values support the contrary proposition. Cells filled with (---) mean that none of the five terms (one contemporaneous plus the first four lags) of the relative CS index is survived to the testing down procedure. Clearly, in these cases the best model turns out to be the baseline one.

The picture emerging from the empirical experiments indicates some intriguing fact. The upper rows of Table 2 show that the total consumption growth depends especially on present and personal economic circumstances, while questions about general conditions and subjective expectations do not show bold values, that is do not have extra information content. The out-of-sample incremental ability of qualitative indices is quite low. As already emphasised by literature, during extraordinary periods as the "tormented" 1992-2003 one in Italy, the emotivity of spending decisions increases. For instance, the variance explained by the model augmented with the General and the Personal indices (+Gen. +Pers.) amount to 76%, whereas the corresponding statistic for the baseline model is 70%. In other sub-samples the difference between these statistics is much smaller. Looking at the "Durable Goods" test, things change. The Overall and the Present indices result orthogonal, while models using other indicators show better performances even if computed throughout different samples. Also, there is evidence of a greater forecasting power as indicated by the RMSPE bold values. The experiment regarding the consumption for non-durable goods suggests that it has emotional content especially in the long term, the most part of bold values being in the full sample estimation. It can be due to the low volatility of this item as compared to the much more schizoid behaviour of household sentiment. Another result is that couples of indices perform better than single indices. In particular, the General and the Personal indices appear to be good for in sample analyses, while the pair Present and Future presents better out of sample statistics. Turning the attention to the expenses for services, findings suggest that the answers about the expectations and the personal feelings of participants in the economy provide the most useful insights on the functioning of this aggregate. As pointed out by the sub sample analyses, adding qualitative information may be very useful in particular periods. The *ad hoc* trials referring to peculiar collections of disbursements, Happiness and Illness, support the intuition that some kind of outlay is more emotional driven than others. With the exemption of the dynamic forecasting exercises, the CS-augmented models outperform the relative baseline in the most part of the empirical designs. On the other hand, results for Illness show that in no case the subjective indicators survive the testing down procedure[7].

By looking vertically downward table 2 we can note that analysing and drawing conclusions on the aggregate information stemming from surveys could be misleading or, more properly, that it could not be the best way to exploit the subjective feelings collected among economic agents. The column "+ Overall" has the less number of bold values which are aimed to indicate the situations where the CS-augmented model outdoes the corresponding baseline one. Similarly, poor performances are realised by the General index which, on the other hand, is the best choice in order to perform dynamic forecasting exercises about non durable goods. Good results are achieved by using the Personal index, in particular when it is applied to services and to Happiness goods or when it is coupled to the General index. The "+Gen.+Pers." pair beats the baseline model seventeen out of twenty five trials, whereas the model augmented by Present and Future turns out to be the best as refer to the out-of-sample dynamic forecast of the consumption for durable and non durable goods. Finally, the out-of-sample experiments performed over the last four years (2000-2003) show better results when using some sub indices. Thus, even during the recent puzzling behaviour of the Italian climate measures, languishing in a very low level as opposed to a relatively sustained growth in the total consumption, a disaggregated analysis is still able to uncover the predictive ability of some indicators.

[7] The low predictive power of this baseline model could be link to the absence of demographic variables which logically should be inserted. However, as explained in the text (see section 3), what is sufficient here is obtaining empirically admissible and coherent models which can be verified by the residuals analysis.

6. Concluding remarks

This paper presented new evidence on the extra information content of surveys of people in households beyond that already contained in quantitative variables. In doing that, a two-step procedure to determine the in-sample and out-of-sample informative power of consumer confidence is performed. First, starting from a general unrestricted ADL model with standard economic regressors, parsimonious consumption equations are found, estimated and used as a forecasting tool. Second, their adjusted R-squared and root mean square errors are compared with those of corresponding models which differ from them only because of the presence of the confidence indicator. The main novelty of this work is that it focuses on the relationships between subjective and objective information at a less aggregate level than that usually analysed. In particular, the overall sentiment index is divided into four sub-indices related to the opinion about the i) current, ii) future, iii) general, and iv) personal situation. In turn, the total private consumption is divided in five items, namely i) consumption for durable goods, ii) consumption for non durable goods, iii) services, iv) consumption for "happiness", v) consumption for "illness".

The main conclusions of the paper can be resumed as follows. Surveys data seems to provide autonomous information. No single index has imposed itself as being clearly superior to the others. The worst results are realised by the most aggregate measure, while much better performances can be obtained by associating a particular sub index to a particular consumption aggregate. The emotional content of spending decisions is different depending on the nature of the goods/services under scrutiny. It is different over time as well, especially if there happen to be extraordinary episodes. All in all, these findings may help to explain why the literature, analysing more aggregate quantitative and qualitative information, shows mixed results. Finally, a word of caution. It may be that a different weighting system could increase the information content of the overall index. But this is beyond the scope of this paper and it is relegated in the agenda.

References

Akaike H. (1973), Information theory and an extension of the maximum likelihood principle. In Petrov, B.N. and Saki F.L. (eds.) Second International Symposium of Information theory. Budapest.

Batchelor R. and P. Dua (1998), Improving Macro-economic Forecasts: The Role of Consumer Confidence, *International Journal of Forecasting*, vol. 14, n. 1, 71-81.

Berg L. and R. Bergstrom (1996), Consumer Confidence and Consumption in Sweden, *Uppsala University Working Paper*, 7.

Bram, J. and S. Ludvigson (1998), Does Consumer Confidence Forecast Household Expenditure? A Sentiment Index Horse Race, *Federal Reserve Bank of New York Economic Policy Review*, 4, June, 59–78.

Campbell, J.Y. and G.N. Mankiw (1989), Consumption, Income, and Interest Rates: Re-interpreting the Time Series Evidence, In *NBER Macroeconomics Annual*, O.J. Blanchard and S. Fischer (eds.), Cambridge: MIT Press.

Campbell, J.Y. and G.N. Mankiw (1990), Permanent Income, Current Income and Consumption, *Journal of Business and Economic Statistics*, 8, July, 265–79.

Carroll C.D., J.C. Fuhrer and D.W. Wilcox (1994), Does Consumer Sentiment Forecast Household Spending? If So, Why?, *American Economic Review*, 84, December, 1397–1408.

Davidson J.E.H., Hendry, D.F. Srba, F. and Yeo, J.S. (1978), Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom, *Economic Journal*, 88, 661–692.

Golinelli R. and Parigi G. (2003), What is this thing called confidence? A comparative analysis of consumer confidence indices in eight major countries, *Temi di Discussione*, Banca d'Italia, 484, September.

Engle R.F., D.F. Hendry, and Richard J.F. (1983), Exogeneity, *Econometrica,* 51, 277-304.

Hall R.E. (1978), The Stochastic Implications of the Life Cycle–Permanent Income Hypothesis: Theory and Evidence, *Journal of Political Economy*, 86, December, 971–87.

Hendry D.F. (1995), *Dynamic Econometrics*. Oxford: Oxford University Press.

Howrey E.P. (2001), The Predictive Power of the Index of Consumer Sentiment, *Brookings Papers on Economic Activity*, 1, 175-216.

Katona G. (1951), *Psychological Analysis of Economic Behaviour*, McGraw Hill, New York.

Katona G. (1960), *The Powerful Consumer: Psychological Studies of American Economy*, McGraw Hill, New York.

Leeper E. M (1992), Consumer Attitudes: King for a Day, *Federal Reserve Bank of Atlanta Economic Review,* July, 1–15.

Martelli B. (1998), Le Inchieste Congiunturali dell'ISCO: aspetti metodologici, *Rassegna di Lavori dell'ISCO*, Vol. XV, n. 3, Rome.

Matsusaka J. and A. Sbordone (1995), Consumer Confidence and Economic Fluctuations, *Economic Inquiry*, 33, 2, 296-318.

Muellbauer J and R Lattimore (1995), *The Consumption Function: A Theoretical and Empirical Overview*, in Handbook of Applied Econometrics, M.H. Pesaran and M. Wickens (eds.), Blackwells.

Mehra Y.P. and E.W. Martin (2003), Why Does Consumer Sentiment Predict Household Spending?, *Federal Reserve Bank of Richmond Economic Quarterly*, 89, 4, 51-67.

Nelson C.R. (1987), A Reappraisal of Recent Tests of the Permanent Income Hypothesis, *Journal of Political Economy*, 95, June, 641–46.

Schwarz G. (1978), Estimating the dimension of a model, *Annals of Statistics*, 6, 461–464.

Wickens M.R. and Breusch T.S. (1988), Dynamic specification, the long run and the estimation of transformed regression models, *Economic Journal*, 98, 189-205.