
Consumer Sentiment and Economic Activity: A Cross Country Comparison

Roberto Golinelli* and Giuseppe Parigi**

Abstract

The objective of this article is to reassess the validity of the consumer confidence (or sentiment) indices in anticipating the evolution of economic activity by considering a fairly high number of countries across the world (i.e. France, Germany, Italy, UK, USA, Japan, Canada and Australia) over a period of about thirty years, from the beginning of the seventies till the end of 2002 (quarterly data). To our knowledge this is the first attempt to analyse the consumer confidence index for several countries over such a long period of time. We model the CSI-output relationship in a cointegrated vector autoregression (VAR) framework, by considering a common set of variables for all countries. Our findings suggest that: (a) what appears to be the main driving forces of consumer confidence cannot be simply summarised on the basis of the most common and used macroeconomic variables; (b) consumer confidence indices have some ability to forecast the evolution of economic activity, provided that both their coincident nature is taken into account and that a number of data-coherent parameter restrictions are imposed in the VAR specifications. These results appear to be fairly robust to several checks we have been able to perform given the availability of a large number of observations. In particular, the forecasting ability of consumer confidence indices is assessed with both in-sample and out-of-sample analyses.

Key Words: Consumer sentiment; GDP indicator; In- and out-of-sample forecasting ability

JEL Classification: C32; C51; C52; E32

* University of Bologna, Department of Economics, Strada Maggiore 45, 40125, Bologna, Italy, golinell@spbo.unibo.it

** Bank of Italy, Research Department, via Nazionale 91, 00184 Rome, Italy, giuseppe.parigi@bancaditalia.it

Résumé

L'objet de cet article est de réexaminer l'intérêt des indices de confiance (ou du sentiment) des consommateurs pour anticiper l'évolution de l'activité économique en passant en revue un nombre relativement élevé de pays du monde entier (à savoir la France, l'Allemagne, l'Italie, le Royaume Uni, les Etats-Unis, le Japon, le Canada et l'Australie) sur une période d'une trentaine d'années, s'étendant du début des années soixante-dix à la fin de 2002 (données trimestrielles). A notre connaissance, c'est la première fois que l'on a tenté d'analyser l'indice de confiance des consommateurs dans plusieurs pays sur une aussi longue période. Nous avons modélisé la relation indice de confiance des consommateurs/production dans un cadre d'auto-régression vectorielle co-intégrée, en examinant une série commune de variables pour tous les pays. Il ressort de nos constatations que : (a) ce qui paraît être les principaux éléments moteurs de la confiance des consommateurs ne peut être simplement ramené aux variables macro-économiques les plus communes et les plus fréquemment utilisées ; (b) les indices de confiance des consommateurs peuvent dans une certaine mesure prévoir l'évolution de l'activité économique, sous réserve que l'on tienne compte de leur nature coïncidente et que l'on impose un certain nombre de restrictions aux paramètres de manière cohérente au niveau des données dans les spécifications de l'auto-régression vectorielle. Ces résultats ont, semble-t-il, bien résisté à plusieurs vérifications que nous avons pu effectuer grâce à l'existence d'un grand nombre d'observations. En particulier, la capacité prévisionnelle des indices de la confiance des consommateurs est évaluée au moyen d'analyses hors échantillon et en échantillon.

1 Introduction¹

Ten years ago, in coincidence with the fortieth birthday of the Michigan index of consumer sentiment, Curtin wrote: „*Consumer sentiment is now the most closely watched and intensely debated indicator of future economic trends*” (Curtin, 1992, p. 22). On its fiftieth birthday this statement is still all the more valid and is the central topic of the debate on the usefulness of sentiment indices. The sentiment index appeared on the economic scene almost by chance, as part of a survey devised by Katona at the Survey Research Centre (SRC) of the Michigan University to investigate the determinants of the financial decisions of households (originally the survey was funded by the Federal Reserve Board).

In the second half of the fifties, the Board of Governors appointed a committee – the so called Smithies committee for the Chair of A. Smithies – to evaluate the validity of the survey data in anticipating consumption behaviour. The broadly negative conclusions of the committee's final report (see FED, 1955) were deeply contested by Katona and his associates. The subsequent debate (see Tobin, 1959; National Bureau of Economic Research, 1960 which contains a selection of the papers presented at a conference on anticipation data organised by the NBER in 1957; and Juster, 1964) strengthened the suspicions that most of leading economists held about the usefulness of the survey. Notwithstanding the firmly negative opinion of mainstream economists on the usefulness of the survey, the SRC continued to collect data and publish the consumer sentiment index over the years.

However, the debate is still open. Indeed, these indices are very popular, being currently reported in the media and commented by economic analysts. This is essentially due to the fact that they are released very promptly and are the only source of information on the evolution of the economy for some time (a preliminary version of the Michigan consumer sentiment index is published during the reference month). Moreover, their earlier availability compared to standard macroeconomic variables may be quite useful for any forecasting exercise. In this context, the relationship between consumer sentiment and general economic activity (or output fluctuations) should be further explored so as to highlight the possible presence of coincident and/or leading links.

From the theoretical point of view given the rational expectations hypothesis, the consumer sentiment index has not to have additional information if it is simply defined as an expected value of macroeconomic variables. However, there is the empirical evidence for which the results are mixed (as surveyed in Section 2).

¹ Paper presented at the CIDE seminar at Milan Catholic University on January 23rd, 2003. We are grateful to Gerard Adams, Filippo Altissimo, Guja Bacchilega, Vieri Ceriani, Clive Granger, Marco Magnani, Paolo Paruolo, Lucio Picci, Federico Signorini, Bruno Sitzia, Stefano Siviero, Daniele Terlizze, Ignazio Visco, Kenneth West, the Editor and two anonymous referees of the Journal for their helpful comments, and to Piera Appoggi for brilliant research assistance. The usual caveats apply. The views contained here are those of the authors only and do not necessarily reflect those of the institutions for which they work. Financial support from MIUR (Roberto Golinelli) is gratefully acknowledged.

The objective of this article is to reassess the empirical validity of the consumer sentiment indices in anticipating the evolution of economic activity by considering eight countries – Australia, Canada, France, Germany, Italy, Japan, UK and USA – over a period of quarterly data of about thirty years, from the beginning of the 1970s to the first quarter of 2002. To our knowledge, this paper is the first to analyse systematically the relationships between consumer sentiment, output fluctuations and other macroeconomic variables over many countries and over such a long period of time by modelling them simultaneously in the cointegrated vector autoregression (VAR) framework. In this way, we appropriately handle the simultaneity concept between non stationary variables, and avoid the limits and drawbacks of the single-equation approach used in most of the literature.

The preliminary step of our analysis is the definition of the set of the variables of interest. Since the potentially large number of variables that may influence both consumer sentiment and output complicates the analysis, a number of exclusion restrictions are tested. This empirical analysis is presented in Sections 2 and 3. The results can be summarized as follows: (1) the consumer sentiment to output relationship presents significant differences across countries; (2) consumer sentiment is related to a set of macroeconomic variables that may change over time.

These two findings support the view that consumer confidence is a more general concept that cannot be summarized only on the basis of some macroeconomic variables. Differences across countries and over time suggest that other factors (*e.g.* psychological) may be at work. Actually, had we found a satisfactory and comprehensive representation of consumer sentiment, its usefulness would have been confined only to the assessment of the cycle with no leading feature for output evolution.

The analysis of the forecasting ability of consumer sentiment is performed in Section 4, where we show that the index has power over that of other macroeconomic variables. However, a crucial element - ignored in the literature - is its contemporaneous link with output fluctuations, particularly relevant for the USA. We also find mixed evidence for the role of consumer sentiment during exceptional periods of the cycle. Section 5 draws some conclusion.

2 The Setting of the Multivariate Systems by Country

Log-levels of the quarterly consumer sentiment or confidence index (CSI) of Australia, Canada, the four biggest European countries, Japan and the USA are plotted in Figure 1 for the period 1970 Q1 - 2002 Q1 (see the Appendix for data sources and definitions). In order to check for robustness, the USA results in the paper will be determined by using both the Michigan and the Conference Board indices.

From the analysis of the graphs in Figure 1 it emerges that the USA indices display the widest fluctuations, with a high degree of synchronization with the output cycle. For Australia, Canada, Germany, France and the UK the cyclical path is much smoother, while for Japan and Italy it presents only mild fluctuations. In general, all CSI series appear to be characterized by

persistence, slow mean reversion and occasional breaks, hence are well represented by stochastic trend processes. These features, confirmed by a number of not reported unit root tests, are quite typical of many time series (such as inflation, unemployment, and interest rates) independently of the economic meaning of non-stationarity (see Brunello *et al.*, 2000, p. 158). In any case, we can safely dismiss the hypothesis that these features are a mechanical result of the methodologies used to calculate the indices, as they are fairly heterogeneous across countries. Although the Michigan survey has been the model for other surveys worldwide, each national institute has adapted the set of questions to national peculiarities, thus providing different measures of sentiment.

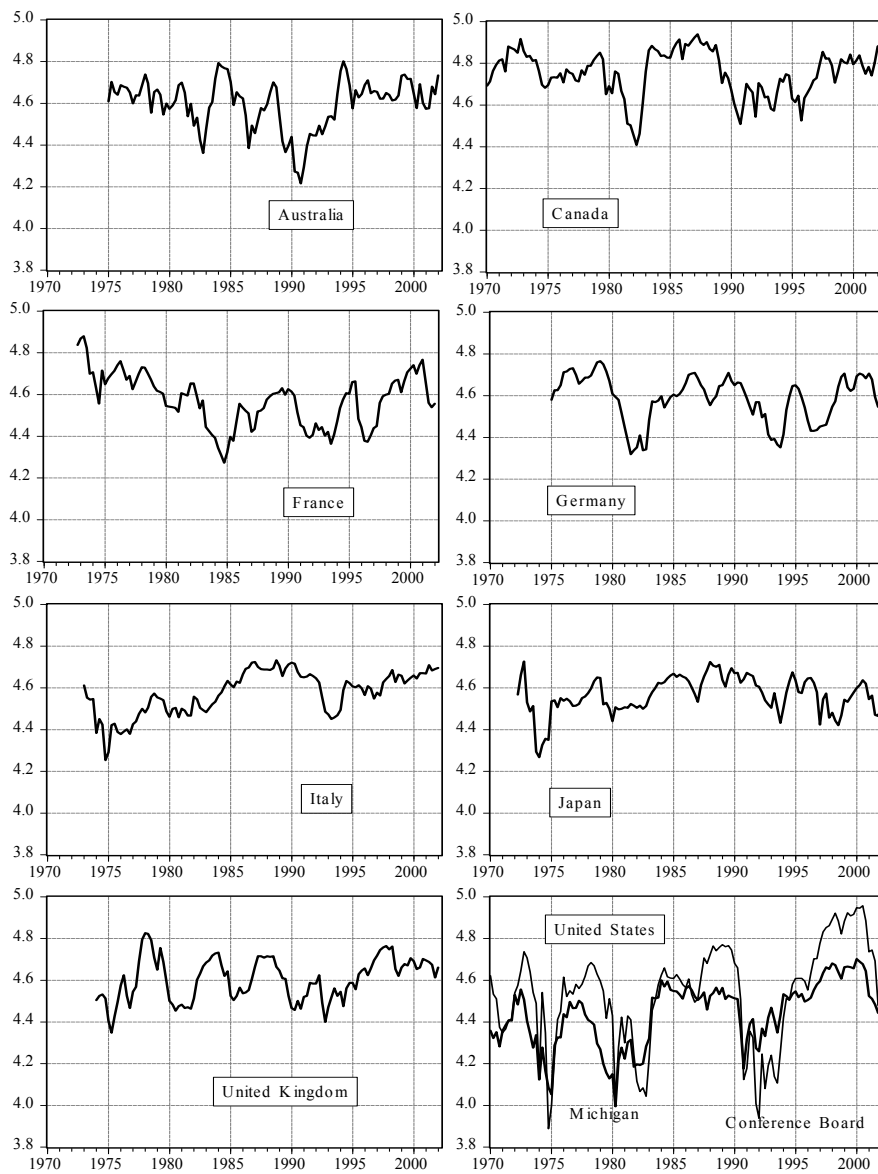
Most of the empirical analyses aimed at establishing whether CSIs had additional information content with respect to traditional macroeconomic variables (*e.g.* inflation, unemployment and output) obtained mixed results: in some cases it was shown that these indices could maintain an autonomous role in forecasting and as explanatory variables in the consumption function (see Mueller, 1963; Adams, 1964; Suits and Sparks, 1965; Fair, 1971a and 1971b; Adams and Klein, 1972); in others, that they could be seen as nothing more than a synthesis of macroeconomic indicators (see Friend and Adams, 1964; Adams and Green, 1965; Hymans, 1970; Juster and Wachtel, 1972a and 1972b; Shapiro, 1972; McNeil, 1974; Lovell, 1975).

The issue of the role of the CSI is still discussed, although the prevailing opinion now seems to be that it may help predict the evolution of economic activity (see Garner, 1991; Fuhrer, 1993; Carrol *et al.*, 1994; Kumar *et al.*, 1995; Matsusaka and Sbordone, 1995; Eppright *et al.*, 1998, Bram and Ludvigson, 1998). In particular, it has been found that its forecasting power tends to be completely offset by other indicators during ordinary times, while it increases notably in the presence of unusual events (see Mishkin, 1978; Throop, 1992; Leeper, 1992; Fuhrer, 1993).

This confirms the original suggestion of Katona (1977) that the CSI is influenced by psychological factors which become particularly important in coincidence with special events, when people are more likely to change their attitude. More recently, Howrey (2001) has shown that the Michigan CSI is characterised by extra forecasting power with respect to other indicators, not limited to exceptional periods (see also Garner, 2002). In any case, it should be stressed that even if the index were just a synthesis of traditional indicators, it could nevertheless maintain, given its timeliness, a great importance for short-term analysis.

The majority of the analyses about the role of the CSI in Europe seem to support the view that it has some autonomous forecasting power (see Van den Abeele, 1983; Praet and Vuchelen, 1984, 1988 and 1989; Praet, 1985; Strümpel and Ziegler, 1988; Batchelor and Dua, 1992; Djerf and Takala, 1997, Nahuis, 2000). In other countries around the world the results are not clear-cut and tend to play down the importance of the CSI (see Santero and Westerlund, 1996, for a general overview of OECD countries; Fan and Wong, 1998, for Hong Kong; Roberts and Simon, 2001, for Australia).

Figure 1 Consumer sentiment indices by country ¹⁾



1) Indexes 1995 = 100 in log-levels

Almost all the analyses above use a single-equation approach, with only partial (or without any) account for multivariate links between the CSI and macroeconomic variables. Insufficient attention is paid to the behaviour of the index in different periods and in different countries and no account is given of the effects of the idiosyncratic characteristics of consumers (*i.e.* income, wealth and education levels, see Souleles, 2001; the CSI may also be influenced by elusive factors such as the feeling of happiness, as explained by Graham in the discussion of Howrey's, paper, 2001: „People's answers to questions about their well-being seem to depend mainly on how they are faring economically relative to their neighbours, whether they themselves have had a bad day, or some noteworthy event in the news", p. 214).

We have chosen to model the CSI-output relationship in a VAR framework, by considering a common set of variables for all countries. This enables us to consider the effects of different definitions of the index, different cyclical patterns and some characteristics of households. In the latter case the behaviour (or the psychology) of the households of a country as a whole is assumed to present some specific features, probably related to particular events that occurred far in the past (like the hyperinflation in Germany) or in more recent times (as the sharp deterioration in public finances in Italy and Japan). The relationship between consumer sentiment and macroeconomic variables may also reflect the characteristics of the economic environment, such as the degree of competition of the markets, the flexibility of the economy (especially the labour market), the nature of the welfare state, the strength of political and economic institutions (see Acemoglu et al., 2002).

The list of the variables to be modelled has been defined mostly on the basis of other empirical analyses. More specifically, besides the CSI our information set includes the following variables (measured at quarterly frequency): GDP growth², output gap, the ratio of the public sector borrowing requirement (PSBR) to GDP, the employment rate for the population aged between 15 and 64, foreign confidence indices, interest rates, stock price changes, and the rates of inflation and of unemployment. The sum of the last two, called the „discomfort index", was proposed by Okun (1962) and was originally related to the CSI by Lovell (1975) and Lovell and Tien (2000).³

² Many studies (among others, Bram and Ludvigson, 1998) underline the potential forecasting power of the CSI in the consumption equation. Therefore, bearing in mind this behavioural relationship, it could seem to be advisable to add consumption (together to GDP) in the list of the variables of interest of our VARs. On the other side, the exclusion of consumption expenditure does not prevent our models to appropriately measure the CSI predicting power of GDP, since the consumption omission implies that VAR parameters we estimated below also embody unidentified combinations of the structural parameters of the (unknown) consumption function. On this point, a different approach models each component of the national accounts separately by using appropriate indicators, see Parigi and Schlitzer (1995), and Buffeteau and Mora (2000), while our modelling approach, as shown in Section 4, is closest to that of Baffigi et al. (2004).

³ In the empirical literature on happiness it is shown (see Clarck and Oswald, 1994, and Oswald, 1997) that the relative weight of the unemployment rate is much higher than is implicit in the discomfort index.

We only focus on a subset of variables since most of them can be seen as different measures of basically the same phenomena. In a VAR context, the exclusion of some variables is an inherently difficult task that we have accomplished on the basis of the Toda and Yamamoto (1995) test for Granger non causality (GNC). The quarterly sample periods, slightly different across countries, are fairly long and generally start in the first half of the 1970s. The VAR lag length (minus one, because all the series are at most first-order integrated) was selected on the basis of both the AIC criterion and residual diagnostics.

The second column of Table 1 lists the country-specific subsets of variables that are used in our empirical analysis. Given the tool we used, it is worth noting that our reduction in the VAR dimension is based on the relative explanatory power of alternative variables, and not on structural inferences (as those made by Ludvigson and Steindel, 1999). Hence, to reduce the risk that wrong marginalizations (e.g. induced by collinearity problems) may affect our findings, in Section 3 we also check for the robustness of outcomes to alternative sets of variables.

Table 1 **Country-specific VAR settings**

	Included variables ¹⁾	Excluded variables ¹⁾	GNC ²⁾	Period ³⁾	Lags
Australia	CSI, DLY, DP, R, DSP	CU, N, U	0.084	1975 Q4–2001 Q4	2
Canada	CSI, DLY, DP, R, DSP, CSI [†]	U, EX	0.024	1970 Q1–2002 Q1	3
France	CSI, DLY, DP, R, U	CU, DSP	0.028	1973 Q3–2002 Q1	3
Germany	CSI, DLY, DP, CU, CSI ^{††}	R, DSP	0.028	1975 Q4–2002 Q1	3
Italy	CSI, DLY, DP, N	CU, R, DSP	0.344	1973 Q3–2002 Q1	3
Japan	CSI, DLY, DP, DSP, U	EX, R	0.026	1973 Q1–2002 Q1	3
UK	CSI, DLY, DP, DSP, U	CU, R	0.120	1974 Q3–2002 Q1	2
Michigan	CSI [†] , DLY, DP, DSP, U	CU, R	0.027	1971 Q2–2002 Q1	4
Conference	CSI, DLY, DP, DSP, U	CU, R	0.045	1971 Q2–2002 Q1	4

1) Labels: **CSI**, consumer sentiment index; **DLY**, GDP quarterly growth;
DP, annualized CPI inflation rate; **R**, nominal interest rate; **DSP**, quarterly change in stock prices;
U, unemployment rate; **CU**, output gap; **CSI[†]**, US Michigan index;
CSI^{††}, average of the French and Italian CSIs;
N, employment-population ratio; **EX**, exchange rate against the US dollar.
Data sources and definitions are in the Appendix. After preliminary inspections, all the variables of interest listed above do not exhibit marked seasonal features.

2) P-values of the degrees of freedom adjusted Toda-Yamamoto test.

3) Quarterly data.

3 Within-Country Consumer Confidence Determinants

Since slow CSIs mean reversions may be the outcome of persistent macroeconomic driving forces, stationary relationships between the CSI and its „determinants” may be tested through cointegrated VAR techniques (in this paper the term „determinant” is used without any reference to a particular theoretical model).

Because of the exclusion of some variables and of the conditioning on the stock price changes⁴, VAR models in Table 2 differ from unrestricted VARs in Table 1 in both number of lags and sample periods.

Table 2 Cointegrated VAR analysis by country

	Sample periods ¹⁾	Lag length	Variables in VAR	Exogenous series ²⁾	Cointegration rank ³⁾	Over-identification ⁴⁾
Australia	1976 Q1–2001 Q4	3	4	DSP	2 ^{**}	0.609 [3]
Canada	1970 Q1–2002 Q1	3	5	DSP	2 [*]	0.09 [5] ⁵⁾
France	1973 Q3–2002 Q1	3	5	-	2 ^{***}	0.140 [4]
Germany	1975 Q4–2002 Q1	3	5	-	2 [*]	0.35 [5] ⁵⁾
Italy	1974 Q1–2002 Q1	5	4	-	2 ^{**}	0.674 [2]
Japan	1973 Q2–2002 Q1	4	4	DSP	2 ^{**}	0.755 [2]
United Kingdom	1974 Q4–2002 Q1	3	4	DSP	2 [*]	0.258 [3]
US Michigan	1971 Q1–2002 Q1	3	4	DSP	2 ^{**}	0.104 [3]
US Conference	1971 Q1–2002 Q1	3	4	DSP	2 ^{**}	0.067 [3]

1) Quarterly data.

2) DSP, quarterly change in stock prices.

3) First two statistics of the Johansen (1995) trace test (with the intercept restricted to lie in the cointegration space): *** both statistics are 1% significant; ** the first is 1% and the second is 5% significant; * the first is 1% and the second is 10% significant.

4) Over-identifying restriction p-values (number of restrictions in squared brackets) of the two relationships.

5) Two weak exogeneity restrictions for foreign CSIs are included.

Rank tests and over-identifying restrictions are reported in Table 2. Both the hypothesis of rank two and the parameter restrictions are data admissible and hence support two stationary relationships: one for GDP growth (the simple difference stationary model, except Okun's law in France, Germany and Japan), and the other for CSI levels, whose estimates are reported in Table 3.

⁴ VAR conditioning on stock price changes (DSP) is valid because the weak exogeneity condition is satisfied (see Johansen, 1995, ch. 8): in modelling the multivariate system there is not loss of information from not modelling the DSP determinants. Similar results are in Otoo (1999) for the USA, and in Jansen and Nahuis (2003) for Europe.

Table 3 **Identified CSI level relationships** ¹⁾

	Country-specific explanatory variables ²⁾							Intercept
	DP	DLY	R	U	CU	CSI ⁺ /CSI ^{**}	N	
Australia	2.587 (0.632)	0.0 (-)	-2.587 (0.632)	-	-	-	-	4.706 (0.039)
Canada	12.98 (3.218)	0.0 (-)	-12.98 (3.218)	-	-	1 (-)	-	0.615 (0.141)
France	3.052 (0.492)	0.0 (-)	-3.052 (0.492)	0.0 (-)	-	-	-	4.676 (0.022)
Germany	-2.959 (0.882)	0.0 (-)	-	-	0.0 (-)	0.928 (0.253)	-	0.400 (1.157)
Italy	-0.482 (0.206)	0.0 (-)	-	-	-	-	5.959 (0.854)	2.209 (0.355)
Japan	-1.094 (0.193)	0.0 (-)	-	-1.094 (0.193)	-	-	-	4.656 (0.015)
United Kingdom	0.0 (-)	0.0 (-)	-	-1.986 (0.824)	-	-	-	4.713 (0.059)
US Michigan	-2.318 (0.531)	0.0 (-)	-	-2.318 (0.531)	-	-	-	4.756 (0.063)
US Conference	-2.399 (0.966)	0.0 (-)	-	-2.399 (0.966)	-	-	-	4.730 (0.114)

1) Standard errors in brackets.

2) Labels are in note 1) of Table 1.

- means „not included in the corresponding country-VAR“.

The robustness of previous findings has been checked over different sub-samples. In particular, three 10 year sub-samples have been considered for Australia, Canada, Japan, and the United States: the „troubled 1970s“ (1971-1980), the Reagan era (1980-1988), Bush (1989-1992) and the Clinton era (1993-2001). For the other four countries two sub-samples have been considered: the EMS period (1978-1991) and the last ten years. While for Canada, the UK and the USA we do not detect relevant changes in the estimates, for Australia after 1993 the inflation coefficient shows a negative sign and the role of interest rates loses significance. In the case of continental Europe results appear to be fairly stable over the EMS period 1978-1991, while afterwards the economic discomfort index becomes the main CSI explanatory variable in France and Germany. In the same period, the real

interest rate and the PSBR to GDP ratio become significant in Italy. In 1980-2002, the Japanese CSI is driven by the unemployment rate and the PSBR to GDP ratio.

These changes suggest that the set of CSI explanatory variables may vary over time according to some possibly country-specific events, though in small samples the results are less clear-cut: less information can allow for different explanations of CSI levels.

An important issue concerns the validity of the specification search conducted for each country. To this end we test whether the specification of a country may be accepted by the data of other countries: in this case the country-specific CSI determinants reported above cannot be considered unique (and the corresponding specification is not robust). Since there are seven alternative specifications (the Japanese and American CSI determinants are the same) and eight country-databases, the analysis generates 56 outcomes. In each case we ask the following four questions: (1) Is the number of stationary relationships at least equal to 2? (2) Are the identification restrictions not rejected? (3) Are the CSI determinants significant, and are their signs in line with *a priori* assumptions? (4) Are the short-run CSI dynamics influenced by the estimated CSI-level relationship? The outcomes of this exercise suggest the uniqueness of the relationships reported in Table 3, since at least one answer in four supports the idea that each country-dataset cannot be explained by any of the other six models. More importantly, they do not support the hypothesis that there is a single model valid for all countries. The evolution of the CSI seems to rest on a country-specific and probably (as in the case of Italy and Japan) time-specific set of few determinants. Cyclical and structural factors are at work, reflecting both the historical evolution of the single countries (see, for instance, the relevance of the inflation and unemployment rates for Germany and the UK, respectively) and the particular economic environment (the short-run effects of stock market prices in the English-speaking countries and Japan). Finally, note that these findings are coherent with the different nature of the CSI surveys in each country (details are in the Appendix).

4 The Ability of Consumer Confidence to Predict GDP

Though there is no consensus in the literature, one set of results supports the idea that the CSI is a leading indicator of the evolution of economic activity. Our contribution to this debate is based on both in-sample and out-of-sample analyses.

In-sample analysis tests for the significance of the *direct* and the *indirect* effects of the CSI on GDP growth. Direct effects can in turn be separated into *lagged* and *simultaneous* effects: if lagged effects are significant, the CSI is a leading GDP indicator; if simultaneous effects are significant, the CSI is a coincident GDP indicator. The latter result has some interesting empirical implications: if the CSI simultaneously causes GDP, early GDP estimates may exploit the timeliness of CSI data releases. On the other hand, CSI indirect effects on GDP pass through the influence of CSI shocks on regressors in the GDP equation, except the CSI-regressor.

In all the previous analyses simultaneity among variables was confined to the variance-covariance matrix of the residuals. In principle, it is possible that assessing CSI simultaneous direct effects on GDP growth may alter the characteristics of the CSI-GDP relationship.

Simultaneity issues may be dealt with using conditional VAR models, where short-run dynamics are identified. In our case, the results concerning the significance of the CSI in explaining the simultaneous GDP growth come from the following short-run identification assumptions: (a) in the short run, the weakly exogenous variables have a simultaneous effect on the CSI (*i.e.* consumer sentiment immediately reacts to news about inflation, interest rates, or stock prices); (b) the exogenous variables have no simultaneous effect on GDP (*i.e.* output dynamics are smoothed by the macroeconomic transmission mechanism of shocks); (c) consumer sentiment embodies simultaneous information on GDP changes.

Table 4 **CSI In-sample predictive power of GDP in conditional VAR models**

	Over-identifying restrictions ¹⁾	Conditioning variables ²⁾	Residual correlation ³⁾	Short-run restrictions ⁴⁾	Coincident indicator ⁵⁾	Leading indicator ⁶⁾
Australia	0.733 [7]	DSP, R	0.2095	0.448 [5]	0.818	0.048
Canada	0.098 [9]	DSP, CSI [†]	0.1658	0.608 [9]	0.175	0.032
France	0.384 [7]	-	0.0708	0.971 [6]	0.877	0.064
Germany	0.458 [7]	CSI ^{**} , DP	0.3264	0.873 [8]	0.037	0.005
Italy	0.88 [13]	-	0.0899	0.88 [15]	0.895	0.005
Japan	0.755 [2]	DSP	0.0255	0.365 [3]	0.363	0.000
United Kingdom	0.408 [5]	DSP, DP	0.2479	0.218 [6]	0.001	0.001
US Michigan	0.347 [8]	DSP, DP	0.4196	0.174 [7]	0.007	0.103
US Conference	0.089 [8]	DSP, DP	0.4807	0.106 [7]	0.000	0.006

1) P-values of the restrictions on the parameters of both level relationships and loadings (number of restrictions in squared brackets).

2) See the labels in Table 1.

3) Correlation between the residuals of CSI and GDP equations in the conditional VAR.

4) Short-run over-identifying parameter restriction p-values (number of restrictions in squared brackets).

5) P-value of the null that the simultaneous CSI parameter in the GDP equation is zero.

6) P-value of the null that all lagged CSI parameters in the GDP equation are jointly zero.

Given the VAR results in the previous section, further weak exogeneity restrictions seem to be reasonable for the interest rate in Australia, the inflation rate in Germany, in the United Kingdom and in the USA. The estimates of the new conditional VAR models in Table 4 (the conditioning variables of each model are listed in the third column) do not show major differences with respect to those of the previous section, residuals are well behaved, and the p-values of the over-identifying restriction tests in the second column of Table 4 are larger than (or equal to) those in the last column of Table 2. In some cases, the high and positive

correlation between the GDP and CSI equation residuals (fourth column) supports hypothesis (c) above. Even though assumptions (a)-(c) are not the only way to identify the short-run propagation mechanism of the shocks, the corresponding over-identifying restrictions are largely not rejected (fifth column), and can be further validated in out-of-sample forecasting exercises (see below).

From the results in the sixth column of Table 4, it can be argued that the CSI is a coincident indicator (its short-run information significantly explains simultaneous GDP changes) in Germany, in the UK and in the US. The outcome of the significance tests in the last two columns of Table 4 suggests that the CSI can also be seen as a predictor of GDP evolution. The importance of a proper treatment of simultaneity effects is particularly evident in the case of the US: in the models of the previous section, no significant lagged effect of the CSI for both the Michigan and the Conference Board indices could be found.

The significance tests conducted so far are in-sample tests of the ability of the CSI to forecast GDP, based on the comparison of the predictive content of two nested models subject to estimation uncertainty (the restricted model acts as the benchmark). Inoue and Kilian (2002) argue that such in-sample tests of predictability have more power than out-of-sample tests in some practical cases.

Insignificant parameters of models in Table 4 are restricted to zero and the resulting models are henceforth labelled structural.⁵ The analysis with structural models of GDP responses to a 1% CSI impulse is, at the same time, an in-sample test of GDP predictability and an estimate of both direct and indirect effects of the CSI on GDP. The results in Table 5 show that in all countries the CSI has a significant (based on Monte Carlo standard errors) effect on GDP with some qualifications: (i) for Australia, Canada, France, Italy and Japan GDP responses are significantly different from zero only after one or two periods; (ii) in the remaining countries the simultaneous effect is clearly evident (more for the UK and Germany, less for the USA).

Katona (1977) suggests that, in ordinary times, the importance of the CSI in explaining activity can be hidden by other variables, while it may become significant in exceptional periods characterised by shocks potentially capable of altering consumer behaviour. More recently, Throop (1992) shows that a small structural model for the USA, where the CSI is well explained in normal times by inflation, unemployment and short-term interest rates, breaks down at exceptional points, such as the Gulf War (similar results are found by Ivanova and Lahiri, 2001, but not by Howrey, 2001).

Our models can be used in a simple exercise (in two steps) in order to assess, though very partially, the effectiveness of Katona's suggestion: we checked whether previous

⁵ Our models are labelled as structural (as opposed to VAR models) because they embody a number of not rejected overidentified restrictions in order to interpret the short term simultaneity direction between the variables of interest, and to get rid of estimates not significantly different to zero. However, the parameter estimates of such „structural” models are still not interpretable in the light of explicit behavioural theories.

impulse-response results might depend on the occurrence of particularly strong shocks in the sample. In the first step, we defined six periods as „exceptional“: (1) the first oil shock (from 1973.3 to 1974.2); (2) the second oil shock (from 1979.1 to 1979.3); (3) the stock market crash of October 1987 (from 1987.4 to 1988.1); (4) the Gulf war (from 1990.3 to 1991.1); (5) the EMS crisis (from 1992.3 to 1993.2); (6) the September 11 terrorist attack (from 2001.3 to 2002.1). In the second step, we assessed whether the profile of the impulse-response functions remains unchanged when the periods above are excluded from the estimation sample by using impulse dummy variables.

Table 5 **Structural models' accumulated GDP responses over T+h to a 1% CSI impulse in T**^{1) 2)}

h (quarters) =	0	1	2	3	4	8	12
Australia	0.00 (-)	0.02 (0.01)	0.03 (0.01)	0.05 (0.02)	0.06 (0.02)	0.09 (0.04)	0.10 (0.05)
Canada	0.00 (-)	0.03 (0.01)	0.08 (0.02)	0.07 (0.02)	0.06 (0.02)	0.04 (0.04)	0.02 (0.05)
France	0.00 (-)	0.00 (-)	0.02 (0.01)	0.02 (0.01)	0.03 (0.01)	0.02 (0.01)	0.02 (0.01)
Germany	0.12 (0.03)	0.17 (0.04)	0.22 (0.05)	0.25 (0.07)	0.28 (0.08)	0.38 (0.14)	0.45 (0.20)
Italy	0.00 (-)	0.08 (0.02)	0.09 (0.03)	0.09 (0.03)	0.12 (0.04)	0.14 (0.06)	0.17 (0.08)
Japan	0.00 (-)	0.10 (0.02)	0.12 (0.02)	0.15 (0.03)	0.18 (0.03)	0.20 (0.06)	0.20 (0.07)
United Kingdom	0.14 (0.03)	0.14 (0.04)	0.18 (0.04)	0.25 (0.05)	0.27 (0.06)	0.36 (0.10)	0.42 (0.14)
US Michigan	0.05 (0.01)	0.07 (0.02)	0.09 (0.02)	0.09 (0.03)	0.09 (0.03)	0.08 (0.04)	0.07 (0.04)
US Conference	0.06 (0.01)	0.07 (0.02)	0.08 (0.02)	0.08 (0.03)	0.08 (0.03)	0.07 (0.04)	0.05 (0.05)

1) % Differences between GDP levels.

2) Standard errors in parentheses (from Monte Carlo experiments with 10,000 replications).

Since the new impulse-response functions remain well inside the confidence intervals of those reported in Table 5, the importance of the role of the CSI cannot be attributed to the effects of exceptional periods only, in line with the findings of Howrey (2001) and Garner (2002).

Of course, in order to better assess this topic, a (symmetric) exercise with models estimated in extraordinary periods had to be made but, unfortunately, the span of extraordinary periods is too short to allow for such an experiment. Therefore, this result must be interpreted with extreme care.

It is widely known (e.g. see Granger, 1990, pp. 3-4) that significant in-sample evidence of predictability does not guarantee significant out-of-sample predictability. The danger with in-sample tests is to detect spurious GDP predictability (i.e. through overfitting) due to the application of particular specification search procedures.

Hence, an alternative way to test for the importance of CSIs in predicting GDP is through an out-of-sample forecasting exercise at different horizons (1, 2 and 4 steps ahead). Table 6 reports the results of GDP forecasts obtained with both structural models (*with* a CSI)⁶, and unrestricted VAR (UVAR) models *without* a CSI. The forecasting performance is evaluated by the comparison of their root mean square error (RMSE) recursively computed from an initial window of 60 quarters. The exogenous variables in the structural models are forecast by simple AR(4) models in differences (nothing substantially changes if either AR(5) on levels or random walk models are used; on this point see, among others, Baffigi *et al.*, 2004).

Overall, the results show that the forecasting power of the CSI extends at most to 2 steps: the 4-steps-ahead forecast RMSEs (as well as that for 3-steps-ahead, not reported) of structural models are generally higher than those obtained with UVAR models without the CSI. We will therefore concentrate on the results for 1- and 2-steps-ahead forecasts. The reduction of the RMSE when the CSI is used is substantial (see the columns 5 and 6) and significant in all countries, with the exception of Japan and the US where, although the results are quantitatively comparable to those of other countries, they are not significant.

Some interesting findings concern the effect of the CSI during the „exceptional“ (shock) periods: the results in columns 8 and 9 seem to support Katona's hypothesis for France, Japan and the USA. The forecasting power of the French CSI (as shown in column 6) is entirely due to its importance in shock periods: the reduction of the RMSE is greater than 20% and 5% significant. For the USA, the reduction is quantitatively large (and significant) especially for the Michigan index. This is more so for the 1-step-ahead exercise when the CSI is assumed to be known over the forecasting horizon: the RMSE for both the Conference and the Michigan index is halved. This is a further indication of the coincident nature of the US index. In other cases, the very small number of forecast periods may affect the results; for example, the reduction of the RMSE in Canada is very substantial, but not significant.

In a forecasting context, Clements and Hendry (2002, p. 321) suggest that, in practice, a forecaster has three main modelling alternatives: (a) using a general unrestricted model; (b) imposing *a priori* restrictions; (c) following a sequential testing procedure. In our case, the UVAR model without the CSI is a model with *a priori* restrictions, while the structural model is

⁶ The FIML estimated parameters, main regression diagnostics, and residual tests by equation and at system level for each structural country-model are not reported here, but are downloadable as a pdf file from the following link: http://www.dse.unibo.it/golinelli/research/GoPa_structural.pdf

a model obtained from the general-to-specific strategy. In order to complete the picture drawn by Clements' and Hendry's paper, we can also define the UVAR model with the CSI as the general unrestricted model. The results from the comparison of the RMSEs of structural and general models (both including the CSI) are qualitatively similar to those above, and the lower RMSEs of structural models can be interpreted as an indication of the importance of reducing the huge number of parameters of the UVAR model.

Table 6 **RMSE of GDP out-of-sample forecasts from alternative models** ¹⁾

	Structural model RMSE ²⁾			Ratio ³⁾ of structural on UVAR without CSI ⁴⁾			Ratio of structural on UVAR without CSI models: shock periods ⁵⁾	
	1-step	2-steps	4-steps	1-step	2-steps	4-steps	1-step ⁶⁾	2-steps
Australia	0.64	0.93	1.47	0.79 ⁺	0.76 ⁺	0.84	0.97	1.12
Canada	0.63	1.02	1.85	0.91	0.82 ⁺	1.27	0.76	0.77
France	0.45	0.73	1.42	1.02	0.90 ^{**}	1.66	1.02	0.78 ^{**}
Germany	0.80	1.29	2.20	0.80 ^{**}	0.87	1.09	0.79 (0.88)	0.82
Italy	0.59	0.96	1.95	0.92 ⁺	0.96	1.03	0.96	0.75
Japan	1.01	1.28	1.52	0.92	1.01	1.15	0.51 ^{**}	0.92
United Kingdom	0.47	0.76	1.33	0.80 ^{**}	0.82 ⁺	1.45	0.85 (1.00)	1.15
US Michigan	0.60	0.92	1.46	0.86	0.86	0.92	0.56 ⁺ (0.48 ⁺)	0.62
US Conference	0.63	0.97	1.68	0.91	0.90	1.07	0.66 (0.49 ⁺)	0.60

1) Forecast horizons: 1990.1-2002.1 for Australia and Germany; 1984.1-2002.1 for Canada; 1988.1-2002.1 for France; 1989.1-2002.1 for Italy and the UK; 1987.3-2002.1 for Japan; 1985.3-2002.1 for the USA.

2) Conditional VAR models with over-identification restrictions also used in impulse-response exercises of Table 5.

3) ⁺, ^{**} Mean that the ratio is 10% and 5% significantly lower than 1 according to the Harvey et al. (1997) testing procedure.

4) The list of the variables in each UVAR model is that of the second column of Table 1, except the CSI; UVAR lag length (on the basis of the AIC criterion.): 4 for Australia, 2 for Canada, France, and the UK, 3 for Germany and the USA, 4 for Japan, and 5 for Italy.

5) The shock periods are: 1987.4-1988.1 (the stock market crash of October 1987) for Canada, Japan and the USA only; 1990.3-1991.1 (the Gulf war); 1992.3-1993.2 (the EMS crisis) for the European countries only; 2001.3-2002.1 (September 11 attack).

6) In parentheses the results when the CSI is assumed to be known over the forecasting period.

The outcome of both in-sample and out-of-sample analysis confirms the predicting power of the CSI, as a leading and as a coincident GDP indicator depending on the country: in Australia, Canada and Europe the index seems to be characterized by leading properties of one-two quarters, while for Japan and the USA the indices appear to be particularly useful (and coincident for the US case) in shock periods.

5 Concluding Remarks

At the end of the fifties, extensive tests on the validity of the SRC survey on consumer intentions and attitudes were basically conducted at the microeconomic (individual) level, by comparing respondents' answers with their subsequent economic behaviour. This was deemed relevant as well as necessary by Tobin who argued: "the necessity of testing at the individual level the predictive value of attitudes and intentions follows inexorably from the inadequacy of any other kind of test." (Tobin, 1959, p. 2). Indeed, as Okun had shown in NBER (1960), results of tests at the aggregate level were at best inconclusive given the availability of very few observations. Since then, a number of papers have tackled the issue of the predictive value of the CSI for economic activity, with often conflicting results. Indeed, most analyses have been characterised by a somewhat restrictive approach, either limited to one single country or to a specific time period (or both). The evidence presented in this paper is the first attempt to our knowledge where the link between the CSI and output is evaluated by simultaneously taking into account several countries (those belonging to the G7 group and Australia) over a fairly long time period (from the beginning of the 1970s to the beginning of 2002).

By exploiting different statistical techniques (in-sample Granger non-causality tests and impulse-response analyses, and out-of-sample forecasting ability) we have highlighted some properties of the CSI with some original findings: a) CSIs have a significant and quantitatively relevant effect on the evolution of GDP; b) CSIs lead GDP independently from other macroeconomic variables; c) in some countries the leading property of the CSIs emerges only after taking into account their simultaneous link with GDP. This helps to explain the often contradictory results found in the literature: for the US, ignoring the simultaneity effect might lead to the mistaken conclusion that the CSI has no forecasting power (especially if the Michigan index is used).

The findings of our paper seem to confirm the view of Katona that confidence indices are influenced by economic as well as other factors. In particular, the CSI does not seem to be correlated to a single set of variables across countries: „classical” variables, such as the rates of inflation and unemployment, are relevant for some countries but not for others. The set of CSI determinants may also change over time. Consider the results for Italy: until the end of the 1980s the evolution of the CSI was driven by inflation and labour market variables. In the last part of the sample the deteriorating state of public finances (the debt to GDP ratio rose over 100% at the beginning of the 1990s) became one of the determinants of the CSI at the expense of the inflation rate (which in the meantime had slowed down to below 10%). A similar description applies to the Japanese case, where the change of the ratio of public debt to GDP becomes significant over the 1980s (again at the expense of the inflation rate). This could imply that the situation of public finances may have an increasing influence of on households' confidence in those countries (such as France and Germany) where there is an intense political debate on the reform of the welfare state. Finally, the evolutionary nature of confidence does not preclude that differences across countries disappear as economies tend to become more similar: the results for France and Germany in the 1990s do in fact suggest

that convergence may matter also for concepts like consumer confidence. These results justify the use of the CSI in the construction of reliable indicators (coincident as well as leading) of the business cycle.

In this context an important issue for further research is the economic interpretation of the link between CSIs and economic activity: some researchers have suggested that CSIs may capture the effects of uncertainty, herd behaviour and/or psychological factors (see Fuhrer, 1988; Acemoglu and Scott, 1994; Matsusaka and Sbordone, 1995; Belessiodis, 1996; Berg and Bergstrom, 1996; Locarno and Parigi, 1997; Parigi and Schlitzer, 1997; Madsen and McAleer, 2000). Another important issue, related to the previous one, is the way the CSI and similar indices are computed, including the procedures of data collection, the kind of questions asked to consumers and the aggregation methodologies. More specifically, a better understanding of consumer behaviour should be obtained by eliciting from the surveys an estimate of the subjective probability distribution of respondents (see Dominitz and Manski, 2003 and Manski, 2003; for an interesting application to firm investment decisions, see Guiso and Parigi, 1999).

References

- Acemoglu, D. and A. Scott. 1994. „Consumer Confidence and Rational Expectations: Are Agents' Beliefs Consistent with the Theory?,” *Economic Journal* 104:422, pp. 1-19.
- Acemoglu, D., S. Johnson, J. Robinson and Y. Thaicharoen. 2002. „Institutional Causes, Macroeconomic Symptoms: Volatility, Crises and Growth,” NBER Working Paper 9124.
- Adams, F. G. 1964. „Consumer's Attitude to Buying Plans and Purchases of Durable Goods: a Principal Components Time Series Approach,” *Review of Economic and Statistics* 46:4, pp. 347-55.
- Adams, F. G. and E. W. Green. 1965. „Explaining and Predicting Aggregate Consumer Attitudes,” *International Economic Review* 6:3, pp. 275-93.
- Adams, F. G. and L. R. Klein. 1972. „Anticipatory Variables in Macroeconomic Models” in *Human Behaviour in Economic Affairs: Essays in Honour of George Katona*. B. Strümpel, J. N. Morgan and E. Zahn, eds. New York, Elsevier Scientific Publishing, pp. 299-320.
- Baffigi, A., R. Golinelli and G. Parigi. 2004. „Bridge Models to Forecast the Euro Area GDP,” *International Journal of Forecasting*, forthcoming.
- Batchelor, R. and P. Dua. 1992. „Survey Expectations in Time Series Consumption Function,” *Review of Economics and Statistics* 74:4, pp. 598-606.
- Belessiodis, T. 1996. „Consumer Confidence and Consumer Spending in France,” *European Commission Economic Papers* 116.
- 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,” *Economic Policy Review* 4:2, Federal Reserve Bank of New York, pp. 59-78.

- Brunello G., C. Lupi and P. Ordine. 2000. „Regional Disparities and the Italian NAIRU,” *Oxford Economic Papers* 52, pp. 146-77.
- Buffeteau, S. and V. Mora. 2000. „Predicting the National Accounts of the Euro Zone Using Business Surveys,” *INSEE Conjoncture* December, pp. 10-18.
- Carrol, C. D., J. C. Fuhrer and D. W. Wilcox. 1994. „Does Consumer Sentiment Forecast Households’ Spending? If So Why?,” *American Economic Review* 84:5, pp. 1397-408.
- Clarck, A. E. and A. J. Oswald. 1994. „Unhappiness and Unemployment,” *Economic Journal* 104, pp. 648-59.
- Clements, M. P. and D. F. Hendry. 2002. „Modelling Methodology and Forecast Failure,” *Econometrics Journal* 5, pp. 319-44.
- Curtin, R. T. 1992. „The Index of Consumer Sentiment at Forty,” *American Enterprise* 3:3, pp. 18-24.
- Djerf, K. and K. Takala. 1997. „Macroeconomy and Consumer Sentiment: Performance of the Finnish Consumer Barometer after Ten Years,” Bank of Finland Discussion Papers 20.
- Dominitz, J. and C. F. Manski. 2003. „How Should We Measure Consumer Confidence (Sentiment)? Evidence from the Michigan Survey on Consumers,” NBER Working Paper 9926.
- Eppright, D., N. Arguea and W. Huth. 1998. „Aggregate Consumer Expectation Indices as Indicators of Future Consumption Expenditures,” *Journal of Economic Psychology* 19, pp. 215-35.
- European Commission. 1997. „The Joint Harmonised EU Programme of Business and Consumer Surveys,” *Report and Studies* 6.
- Fair, R. C. 1971a. *A Short Run Forecasting Model of the U.S. Economy*. Lexington.
- Fair, R. C. 1971b. „Consumer Sentiment, the Stock Market and Consumption Functions,” Econometric Research Program, Memorandum 119, Princeton University, September.
- Fan, C. S. and P. Wong. 1998. „Does Consumer Sentiment Forecast Consumer Spending? The Hong Kong Case,” *Economics Letters* 58, pp. 77-84.
- Federal Reserve Consultant Committee on Consumer Survey Statistics. 1955. Smithies Committee report in: Reports of the Federal Reserve Consultant Committee on Consumer Survey Statistics, Hearings of the Subcommittee on Economics Statistics of the Joint Committee on the Economic Report, 84th US Congress.
- Friend, J. and F. G. Adams. 1964. „The Predictive Ability of Consumer Attitudes, Stock Prices and Non-Attitudinal Variables,” *Journal of the American Statistical Association* 59:12, pp. 987-1005.
- Fuhrer, J. C. 1988. „On the Information Content of Consumer Survey Expectations,” *Review of Economics and Statistics* 70:1, p. 140-44.
- Fuhrer, J. C. 1993. „What Role Does Consumer Sentiment Play in the U.S. Macroeconomy?,” *New England Economic Review* 1, Federal Reserve Bank of Boston, pp. 32-44.
- Garner, C. A. 1991. „Forecasting Consumer Spending: Should Economists Pay Attention to the Consumer Confidence Surveys?,” *Economic Review*, Federal Reserve Bank of Kansas City, May-June, pp. 57-71.
- Garner, C. A. 2002. „Consumer Confidence After September 11,” *Economic Review*, Federal Reserve Bank of Kansas City, May-June, pp. 5-26.
- Granger, C. W. J. 1990. *Modelling Economic Series. Readings in Econometric Methodology*, Oxford University Press.
- Guiso, L. and G. Parigi. 1999. „Investment and Demand Uncertainty,” *Quarterly Journal of Economics* 114:1, pp. 185-227.
- Harvey, D. I., S. J. Leybourne and P. Newbold. 1997. „Testing the Equality of Prediction Mean Squared Errors,” *International Journal of Forecasting* 13, pp. 281-91.
- Hymans, S. H. 1970. „Consumer Durables Spending: Explanation and Prediction,” *Brookings Papers on Economic Activity* 2, pp.173-99.
- Howrey, E. P. 2001. „The Predictive Power of the Index of Consumer Sentiment,” *Brookings Papers on Economic Activity* 1, pp. 175-216.
- Inoue, A. and L. Kilian. 2002. „In-Sample and Out-of-Sample Tests of Predictability: Which One Should We Use?,” *CEPR Discussion Papers* 3671, London.

- Ivanova, D. and K. Lahiri. 2001. „When Should We Care About Consumer Sentiment? Evidence From Linear and Markov-Switching Models,” *Indian Economic Review* XXXIV, pp.153-69.
- Jansen W. J. and N. J. Nahuis. 2003. „The Stock Market and Consumer Confidence: European Evidence,” *Economics Letters* 79, pp. 89-98.
- Johansen, S. 1995. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press.
- Juster, F. T. 1964. *Anticipations and Survey. An Analysis of Consumer Behavior*. Princeton, Princeton University Press.
- Juster, F. T. and P. Wachtel. 1972a. „Uncertainty, Expectations and Durable Goods Demand,” in *Human Behaviour in Economic Affairs: Essays in Honour of George Katona*. B. Strümpel, J. N. Morgan and E. Zahn, eds. New York, Elsevier Scientific Publishing, pp. 321-45.
- Juster, F. T. and P. Wachtel. 1972b. „Inflation and the Consumer”, *Brookings Papers on Economic Activity* 1, pp. 71-114.
- Katona, G. 1977. *Psychological Economics*, New York, Elsevier.
- Kumar, V., R. P. Leone and J. N. Gaskins. 1995. „Aggregate and Disaggregate Sector Forecasting Using Consumer Confidence Measures”, *International Journal of Forecasting* 11, pp. 361-77.
- Leeper, E. M. 1992. „Consumer Attitudes: King for a Day,” *Economic Review* 77:4, Federal Reserve Bank of Atlanta, pp. 1-15.
- Locarno A. and G. Parigi. 1997. „Clima di fiducia e consumi delle famiglie: movente economico o psicologico?,” *Ricerche quantitative per la politica economica* II, Banca d'Italia, Roma.
- Lovell, M. C. 1975. „Why Was the Consumer Feeling so Bad?”, *Brookings Papers on Economic Activity* 2, pp. 473-79.
- Lovell, M. C. and P. L. Tien. 2000. „Economic Discomfort and Consumer Sentiment”, *Eastern Economic Journal* 26:1, pp. 1-8.
- Ludvigson, S. and C. Steindel. 1999. „How Important Is the Stock Market Effect on Consumption?,” *Economic Review*, Federal Reserve Bank of New York, July, pp. 29-51.
- Madsen, J. and M. McAleer. 2000. „Direct Tests of the Permanent Income Hypothesis under Uncertainty, Inflationary Expectations and Liquidity Constraints”, *Journal of Macroeconomics* 22:2, p. 229-52.
- Manski, C. F. 2003. „Inference on Expectations and Decisions,” mimeo, Department of Economics and Institute for Policy Research, Northwestern University.
- Matsusaka, J. G. and A. M. Sbordone. 1995. „Consumer Confidence and Economic Fluctuations,” *Economic Inquiry* XXXIII:2, pp. 296-318.
- McNeil, T. 1974. „Federal Programs to Measure Consumer Purchase Expectations, 1946-73: a Post-Mortem,” *Journal of Consumer Research* 1, pp. 7-10.
- Mishkin, F. S. 1978. „Consumer Sentiment and Spending on Durable Goods”, *Brookings Papers on Economic Activity* 1, pp. 217-32.
- Mueller, E. 1963. „Ten Years of Consumer Attitude surveys: Their Forecasting Record”, *Journal of American Statistical Association* 58:10, pp. 899-917.
- National Bureau of Economic Research. 1960. *The Quality and Economic Significance of Anticipations Data*, Special Conference Series, Princeton, Princeton University Press.
- Nahuis, N. J. 2000. „Are Survey Indicators Useful for Monitoring Consumption Growth? Evidence from European Countries”, *Monetary and Economic Policy Department Working Paper*, No. 2000-08, De Nederlandsche Bank, Amsterdam.
- Okun, A. M. 1962. „Potential GNP: Its Measurement and Significance,” *Proceedings*, Business and Economics Statistics Section of the American Statistical Association, pp. 98-104.
- Oswald, A. J. 1997. „Happiness and Economic Performance”, *Economic Journal* 107, pp. 1815-31.
- Otoo, M. W. 1999. „Consumer Sentiment and the Stock Market,” *Finance and Economic Discussion Papers*, No. 1999-60, Federal Reserve Board, Washington, D.C.
- Parigi, G. and G. Schlitzer. 1995. „Quarterly Forecasts of the Italian Business Cycle by Means of Monthly Economic Indicators,” *Journal of Forecasting* 14, pp. 117-41.
- Parigi, G. and G. Schlitzer. 1997. „Predicting Consumption of Italian Households by Means of Survey Indicators”, *International Journal of Forecasting* 13, pp. 197-209.

- Praet, P. 1985. „Endogenizing Consumers' Expectations in Four Major EC Countries”, *Journal of Economic Psychology* 6, pp. 255-69.
- Praet, P. and J. Vuchelen. 1984. „The Contribution of EC Consumer Surveys in Forecasting Consumer Expenditures: An Econometric Analysis for Four Major Countries”, *Journal of Economic Psychology* 5, pp. 101-24.
- Praet, P. and J. Vuchelen. 1988. „Exogenous Shocks and Consumer Confidence in Four Major European Countries,” *Applied Economics* 20, pp. 561-67.
- Praet, P. and J. Vuchelen. 1989. „The Contribution of Consumer Confidence indices in Forecasting the Effects of Oil Prices on Private Consumption,” *International Journal of Forecasting* 5:3, pp. 393-97.
- Roberts I. and J. Simon. 2001. „What Do Sentiment Surveys Measure?”, *Research Discussion Paper* 9, Reserve Bank of Australia.
- Santero, T. and N. Westerlund. 1996. „Confidence Indicators and Their Relationship to Changes in Economic Activity”, *OECD Working Paper* 170.
- Shapiro, H. T. 1972. „The Index of Consumer Sentiment and Economic Forecasting: a Reappraisal,” in *Human Behaviour in Economic Affairs: Essays in Honour of George Katona*. B. Strümpel, J. N. Morgan and E. Zahn, eds. New York, Elsevier Scientific Publishing, pp. 373-396.
- Souleles, N. S. 2001. „Consumer Sentiment: Its Rationality and Usefulness in Forecasting Expenditure. Evidence from the Michigan Micro Data”, *NBER Working Paper* 8410.
- Strümpel, B. and R. Ziegler. 1988. „Consumer Sentiment in Europe - Lessons from Fifteen Years,” in *Contributions of Business Cycle Surveys to Empirical Economics*. Oppenländer, K. H. and G. Poser, eds. Avebury, Aldershot, pp. 575-598.
- Suits, D. B. and G. R. Sparks. 1965. „Consumption Regression with Quarterly Data”, in *The Brookings Quarterly Econometric Model of the United States*. J.S. Duesenberry, G. Fromm, L. R. Klein and E. Kulu, eds. Rand-McNally.
- Throop, A. 1992. „Consumer Sentiment: Its Causes and Effects”, *Economic Review* 1, Federal Reserve Bank of San Francisco, pp. 35-59.
- Tobin, J. 1959. „On the Predictive Value of Consumer Intentions and Attitudes”, *Review of Economic and Statistics* 41:1, p. 1-11.
- Toda, H. Y. and T. Yamamoto. 1995. „Statistical Inference in Vector Autoregressions with Possibly Integrated Processes”, *Journal of Econometrics* 66, pp. 225-250.
- Van den Abeele, P. 1983. „The Index of Consumer Sentiment: Predictability and Predictive Power in the EEC,” *Journal of Economic Psychology* 3, pp. 1-17.

Appendix

Data sources and definitions

Label	Definition	Source
CSI	Consumer sentiment (confidence) index	See below
DLY	GDP quarterly growth	$\text{Log}(\text{GDP}/\text{GDP}_{-1})$
GDP	Gross Domestic Product at constant prices	National Accounts (quarterly); national statistical offices.
DP	Annualized quarterly inflation rate	$4 \cdot \log(P/P_{-1})$
P	Consumer price index	OECD, Main Economic Indicators.
R	Short-term interest rate	OECD, Main Economic Indicators
DSP	Quarterly change in stock prices	$\text{Log}(\text{SP}/\text{SP}_{-1})$
SP	Stock price index	OECD, Main Economic Indicators
EX	Exchange rate against the US dollar	OECD, Main Economic Indicators
CU	Rate of capacity utilization	OECD, Main Economic Indicators.
N	Employment-population ratio	OECD, Main Economic Indicators.
U	Unemployment rate	OECD, Main Economic Indicators

For all variables, quarterly values are obtained directly from the main data source.

For the consumer sentiment indices, see below.

Detailed sources of consumer sentiment

Australia. The Westpac-Melbourne Index of Consumer Sentiment is obtained from a monthly telephone survey of 1200-1400 households (the survey began in 1973 and was quarterly until 1976; since then it was conducted every 6 weeks until 1986 and monthly afterwards). The index is calculated by adding 100 to the arithmetic average of the net balances of positive minus negative responses to five questions on: (1) personal financial conditions over the past year; (2) anticipated personal financial conditions over the coming year; (3) anticipated economic conditions over the coming year; (4) anticipated economic conditions over the next five years; (5) whether now is a good or bad time to buy major household items. The final version of the (seasonally adjusted) index is obtained as a ratio to the base-period level. Quarterly values are obtained as simple arithmetic averages of the three months in the quarter.

Canada. For Canada we use the Index of Consumer Attitudes computed from the Conference Board's survey of a random sample of Canadian households. The survey began in 1960 on a quarterly basis (monthly starting in 2001). The index is computed as the arithmetic average of the (seasonally adjusted) net balances of positive minus negative responses to four questions on: 1) the financial condition of the household in the past six

months; 2) the financial condition of the household in the next six months; 3) the short-term (next six months) employment outlook; 4) the convenience of making an outlay for items such as home, car or other major item. The final version of the index is obtained as a ratio to the 1991 level. Quarterly values are obtained as simple arithmetic averages of the three months in the quarter.

France, Germany, Italy and the UK. In all countries monthly telephone surveys (along the lines suggested in European Commission, 1997) of about 2000 households are conducted by INSEE, GFK, ISAE and GFK for France, Germany, Italy and the UK, respectively. In all countries the surveys began in the early 1970s and were conducted three times a year until 1980 for the UK, 1981 for Germany, 1982 for Italy and 1985 for France; the missing summer quarter was obtained as the average of the second and the fourth quarter. For Italy a monthly series from 1973 to 1981 was calculated by Locarno and Parigi (1997).

For France, Germany and the UK we have used the European Commission version of the index, which is computed as the arithmetic average of the (seasonally adjusted) net balances of positive minus negative responses to four 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. The index we use is the ratio to its 1995 level.

For Italy we have used the national version of the index which is based on 9 questions, the previous 4 plus 5 new ones on: (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. 100 is added to the arithmetic average of the balances and the (non-seasonally adjusted) index is computed by ISAE as a ratio to the 1980 level.

For all countries, quarterly values are obtained as simple arithmetic averages of the three months in the quarter.

Japan. The consumer sentiment index has been computed by the Department of Business and Statistics of the Economic and Social Research Institute (ESRI) cabinet office since the beginning of the 1970s. The ESRI conducts a quarterly survey of a stratified random sample of 5040 Japanese households (excluding single-person and foreign households) through mail questionnaires. The households are asked to evaluate what they consider to be the prospects over the next six months for: overall livelihood; income growth; prices; employment; willingness to buy durable goods. A weighted average of answers to each question (improve, +1, improve slightly, +0.75; no change, +0.5; worsen slightly, +0.25; worsen, 0) is used to compile each index (50 is neutral) and the overall consumer sentiment is obtained as the simple average of the five components and then seasonally adjusted. Quarterly values are obtained as simple arithmetic averages of the three months in the quarter.

United States of America. The University of Michigan's Consumer Sentiment Index is computed from the replies to the questions of a monthly telephone survey of at least 500 households conducted by the Survey Research Centre at the University of Michigan. The

index is calculated by adding 100 to the arithmetic average of the net balances of positive minus negative responses to five questions on: (1) personal financial conditions over the past year; (2) anticipated personal financial conditions over the coming year; (3) anticipated economic conditions over the coming year; (4) anticipated economic conditions over the next five years; (5) whether now is a good or bad time to buy major household items. The final version of the (non-seasonally adjusted) index is obtained as a ratio to the base-period level.

The Conference Board Consumer Confidence Index is obtained from a monthly survey, which the Conference Board mails to 5000 households, receiving about 3,500 responses. The index is calculated by adding 100 to the ratio of positive responses to the sum of positive and negative responses to five questions on: (1) present general business conditions in the area; (2) present job availability conditions in the area; (3) anticipated business conditions in the area over the coming six months; (4) anticipated job conditions over the coming six months; (5) anticipated personal financial conditions over the coming six months. The final version of the (seasonally adjusted) index is obtained as a ratio to the base-period level.

In both cases quarterly values are obtained as simple arithmetic averages of the three months in the quarter.