ISTITUTO DI STUDI E ANALISI ECONOMICA

VECTOR-AUTO-REGRESSION APPROACH TO FORECAST ITALIAN IMPORTS

Carmine Pappalardo
ISAE - Piazza Indipendenza, No. 4, 00185 Rome (Italy).
E-mail: c.pappalardo@isae.it
Tel. +39 06 4448 2355

and

Gianfranco Piras
Tor Vergata University, Faculty of Economics
via Columbia, No. 2, 00133 Rome (Italy)

ISAE - Piazza dell'Indipendenza, No. 4, 00185 Rome (Italy).
E-mail: g.piras@isae.it
Tel. +39 06 4448 2347

Working paper No. 42

February 2004

The Series "Documenti di Lavoro" of the *Istituto di Studi e Analisi Economica* – Institute for Studies and Economic Analyses (ISAE) hosts the preliminary results of the research projects carried out within ISAE. The diffusion of the papers is subject to the favourable opinion of an anonymous referee, whom we would like to thank. The opinions expressed are merely the Authors' own and in no way involve the ISAE responsability.
The Series is meant for experts and policy-makers with the aim of submitting proposals and raising suggestions and criticism.

La serie "Documenti di Lavoro" dell'Istituto di Studi e Analisi Economica ospita i risultati preliminari di ricerche predisposte all'interno dell'ISAE. La diffusione delle ricerche è autorizzata previo il parere favorevole di un anonimo esperto della materia che qui si ringrazia. Le opinioni espresse nei "Documenti di Lavoro" riflettono esclusivamente il pensiero degli autori e non impegnano la responsabilità dell'Ente.
La serie è destinata agli esperti ed agli operatori di politica economica, al fine di formulare proposte e suscitare suggerimenti o critiche.

ISAE
Piazza dell'Indipendenza, 4 00185 Roma. Tel. +39-0644482215; e-mail: relazioni.esterne@isae.it

ABSTRACT

Imports represent a relevant component of total economic resources. For the Italian case, they mainly consist of raw materials and intermediate goods.
In this paper, we evaluate several econometric models performing short-horizon forecasts of Italian imports of goods. Year-to-year growth rate of the monthly seasonally unadjusted series is the variable to predict. VAR forecasting ability has been compared to that of a linear univariate benchmark (ARIMA) model. Main forecast diagnostics have been presented. Finally, we perform two types of forecast encompassing tests (Diebold-Mariano, 1995; Fair-Shiller, 1990) for which we present main results.

JEL Classification: C53, C52, C32.

Keywords: Forecasting, VAR model, Import, Forecast evaluation.

We are grateful to an anonymous referee for the helpful comments and to all the participants at the Meeting "Modelli Stocastici e Metodi di Simulazione per Analisi di Dati Dipendenti", University of Molise, Faculty of Economics, 28-29 April 2003, and to all the participants at the ISAE seminar. A particular thank to Claudio Lupi, Giancarlo Bruno, Sergio de Nardis for the useful discussions and suggestions, and thanks to Gianna Foschi for her efficient editorial assistance. The opinions expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of ISAE or its staff.

NON TECHNICAL SUMMARY

Imports represent a relevant component of total economic resources. For the Italian case, they mainly consist of raw materials and intermediate goods. For this reason, imports can be taken as a significant leading indicator of the aggregate business cycle. This feature, though extremely useful in assessing very short-run dynamics of Italian economy, cannot be properly exploited due to the lack of the availability of the statistical information. Quantity indexes, released from ISTAT, are made available about three months late with respect to the reference period. Our aim is to provide very short run forecasts useful to integrate the available information.

In this paper, we evaluate several econometric models performing short-horizon forecasts of Italian imports of goods. Year-to-year growth rate of the monthly seasonally unadjusted series is the variable to predict. VAR forecasting ability has been compared to that of a linear univariate benchmark (ARIMA) model. Main forecast diagnostics have been presented. Finally, we perform two types of forecast encompassing tests (Diebold-Mariano, 1995; Fair-Shiller, 1990) for which we present main results.

LA PREVISIONE DI BREVE TERMINE DELLE IMPORTAZIONI DI BENI IN QUANTITA'

SINTESI

Le importazioni di beni costituiscono una componente particolarmente rilevante nella composizione delle risorse complessive di un paese. In Italia, l'acquisto di beni dall'estero è in prevalenza costituito da materie di base e prodotti semilavorati, utilizzati nelle fasi iniziali e intermedie del processo produttivo. Per questa caratteristica, le importazioni presentano un comportamento anticipatore dell'economia nazionale.
L'esercizio di previsione consiste in un'applicazione della metodologia VAR (Vector Autoregression) alla serie dei volumi mensili di beni importati e alle variabili coincidenti e anticipatrici di tale indicatore. L'analisi della *performance* previsiva è condotta rispetto ad un insieme di previsione ottenute da un modello univariato ARIMA, corretto per tenere conto di valori anomali ed effetti di calendario. I modelli sono stati confrontati utilizzando due diverse metodologie di test di *forecast encompassing* (Diebold-Mariano, 1995; Fair-Shiller, 1990).

Classificazione JEL: C53, C52, C32.

Parole chiave: Previsione, Modelli VAR, Import, Diagnostiche di previsione.

CONTENTS

1. INTRODUCTION

In Italy, external balance represents a very crucial component in determining the size and pattern of real product. As far as external demand is concerned, exports may be considered as a significant share of total national production of goods and services sold abroad. Particularly, during expansive cycles, domestic production of goods of traditional specialization of Italian industry (such as textiles and clotures, furniture, industrial machines) has been mostly sustained by external demand. Also benefiting of significant exchange rate devaluations, Italian exports have largely contributed to the growth of national economy in the last decades.

Raw and intermediate materials, which domestic firms use as inputs in the initial phases of productive process, are mostly bought on international markets. As a consequence, increase of imports (mainly of goods) can also be considered as a signal of a successive rise in output dynamics, driven by a growth of internal and/or external demand components (investments in capital goods, consumption, exports). Moreover, as directly related to the industrial cycles (one of the main sources of aggregate fluctuations), imports represents one of the most significant leading indicators of short-term patterns of Italian economy.

Tough imports share relevant information for short-term investigations, such features cannot be adequately used for business cycle purposes, given the excessive delay with which official statistics are provided. Indicators in volume, which we are much interested in for business cycle purposes, are available about a quarter after the reference month. Hence, a three-step ahead prediction is necessary to achieve a nowcast of the indicator itself. As far as we are concerned, it has not been undertaken a systematic forecast of Italian import of goods on a monthly basis. In this context, short-term forecasting models play a significant role to face the gap between the availability of official statistics and the needs of a timely short-term business cycle analysis.

In this paper, several short-term forecasting models of the Italian imports of goods in volume are presented. In-sample-predictions from a simple univariate model are used as benchmark to evaluate the forecasting performances of alternative specifications of VAR (Vector Autoregression) models. Forecast encompassing tests (Diebold and Mariano, 1995; Fair and Shiller, 1990), based on comparison of the information from ARIMA and VAR predictions, have been carried out as an additional evidence on forecasting accuracy.

The paper is organized as follows: in the next section we present some preliminary analysis of the time series concerned in the specification and forecasting exercises. In particular, departures from the linearity assumption have been tested. Section 3 develops VAR models identification, estimation and forecasting procedures. Section 4 is dedicated to analyze forecast evaluation. Prediction accuracy, evaluated through forecast encompassing tests, has been developed in section 5. Section 6 concludes.

2. PRELIMINARY ANALYSIS

This section provides some descriptive but relevant preliminary information concerning the series involved in the forecasting exercise. For the Italian case, the building of a forecasting model of the annual dynamics of the total import of goods in volume has not been undertaken on a systematic basis. The excessive delay with which the information on the volumes imported is made available, makes difficult the use of such series in common analysis. Recently, a strong effort has been undertaken by the Italian National Statistical Institute (ISTAT) to make quantity indicators on imports (and exports) much more timely. As an additional difficulty, it is not actually available a long time series to be used for prediction purposes. The reference time series, actually updated by ISTAT on a monthly bases, starts from 1996; such series is not immediately comparable with the previous ones due to a deep revision both in statistical methodology both in classification standards adopted.

A long time series has been obtained using all existing information. Actually, 1995 based series are directly provided by ISTAT. Starting from January 1996, data in volume are constructed as chain indexes with the reference basis changing one time a year and covering a large set of goods traded. Also, they take account of the dynamics of imports composition through time, which was considered to be fixed by the old series. Data based on the previous methodology, available over a long time period ranging from 1980 until 1998, consist in a fixed based series (1985=100). The reconstruction of a long time series was firstly based on the comparison of the two indicators (the old and the new one) over a common time span covering the period 1996:01 - 1998:12.

Data before 1996 have been obtained from the old indicator scaled using the parameters from a linear regression of the new and the old series, estimated over the common sample of 36 observations. Such resulting series presents the advantages to be available over a longer time span (starting from 1980) and to be monthly updated on the basis of the indicators in volume currently produced by ISTAT. Such indicators is that used for the identification and estimation of very short-run forecasting econometrics models of the monthly series of total imports in volume. Taking account of the delay with which the official indicator is released, three steps ahead forecasts may be correctly considered as a nowcast of the reference series. Really, *true* forecasts may be obtained through predictions of the official figures over a longer forecasting horizon.

Contrary to other researches (e.g., Bruno and Lupi, 2001) the set of indicators for import is not as rich as those for other variables. At least at this initial stage, roughly coincident variables have been considered too. The variables selected as potential predictors of volumes imported are the following: the general index of industrial production (IPI), the quantity of goods transported by railways (TK), the business survey series of short-term production prospects (TTP), the Italian exports of goods in volume (XWQ). [1]

2.1. Potentially leading indicators

Industrial production is the key indicator for the monitoring of business cycles dynamics. Since early stages of productive process is crucially dependent on raw materials and intermediate goods, increases in production should be relatively anticipated by a growth in imports. IPI has been included considering the strict relationships between production and flows of goods from abroad. Thus, it does not present any particular leading property, as the descriptive analysis in the following sections will prove. The need to take account of production dynamics in forecasting imports led to select two main proxies of industrial production. First of all, following Bruno and Lupi (2001), goods transported by railways proved to be a powerful predictor of industrial production since they largely consist of intermediate and raw materials. TTP is a variable taken from the business survey on manufacturing sectors carried out by ISAE on a monthly basis. It represents industrial operators' opinions on production dynamics in the very short-term (three-four months ahead).[2] Cyclical analysis showed a long lead of production expectations also over imports of goods. Finally, the monthly series of

[1] The industrial production index is released monthly by ISTAT. The business survey series of short-term production prospect is released by ISAE, the Institute for Studies and Economic Analysis. The time series for tons of goods transported by railways are kindly provided by Trenitalia, the Italian State railways company.

[2] Operators are asked to express their opinion on production pattern, according to three modalities: "up", "stable" and "down". The variable has been quantified through the balance approach.

products exported has been considered. It shares the same features of the imports and has been reconstructed in the same way. The assumption is that products sold abroad can be viewed as a share of national production and, hence, are able to activate imports of goods (especially with reference to given specific sectors). In the following sections, univariate characteristics of the above set of predictors will be explored. All series are considered as log transformed while the expression -log((200/(TTP+100)-1) is applied to production expectations to make the series unbounded.

Figure 1 (first five panels) presents the plots of the set of series (to be predict/ predictors) we are dealing with. We provide a description of such series in terms of their long run and medium term stochastic properties. Trend components, which describe the long-run dynamics of the variables, have been assumed to include frequencies exceeding 8 years. Cycle frequencies range from 18 up to 96 months are extracted using Band-Pass filters (see subsection 2.3). Seasonal fluctuations have been previously described according to their evolving pattern. Further, for each variable, the null of seasonal integration has also been tested (see subsection 2.2). Imports show strong seasonal patterns changing over time. In particular, seasonal dynamics present substantial changes at the beginning of 90's, showing a more regular behaviour. They are characterized by significantly increasing fluctuations in the second part of the sample. Imports long run dynamic sharply increases from the mid of 90's, assuming rate of growth largely greater than those observed in the past. The series of quantity exported shows long run dynamics rising along the whole time period. Its seasonal pattern shares features analogous to that of imports, changing over time with large oscillations in the second half of the sample. Industrial production is characterized by a greater regularity of its long term pattern. Oscillations at business cycle frequencies appear to be prevailing respect to the trend, given its moderate rates of growth. Seasonality presents strong and regular fluctuations (with large though in August) so to assume a largely deterministic pattern. A similar pattern shows the railways transport of goods: the series is mostly characterized by medium term (business cycle) frequencies, and possible outliers. A significant different behaviour concern the series of short-term production expectations. The plot shows dominant cyclical patterns; over the sample considered, it appears a lightly growing trend. Such features, will be assessed in more detail in successive sub-sections.

2.2. Stationarity and seasonality

Some relevant stochastic properties of the time series we consider have been evaluated through a unit root test. As we are dealing with raw data, the presence of such roots has been detected both at the zero (regular) both at seasonal frequencies. Empirically, the analysis has been carried out using the test due to Beaulieu and Miron (1993). For each variable, test statistics are obtained running an auxiliary regression, which deterministic part is specified with a constant, a trend, eleven seasonal dummies and lags of the dependent variable up to get white noise residuals. A synthetic evidence is reported in Table 1. Such results reject the presence of a unit root at zero frequency for imports, exports and industrial production variables. Considering TK series, the null of stationarity is accepted both at the regular both at some seasonal frequencies. Short-run production expectations does not show any unit root at the seasonal frequencies. Testing also rejects the null at the zero frequency. Such result does not seems to be coherent with recent prevailing evidences: it could be due to the particular sample extension and to the presence of outliers, which could affect testing. In all cases, the presence of a complete set of unit roots is rejected with strong evidence. In spite of this, the assumption of no-stationarity at frequencies other than the regular is confirmed for all the variables considered in this paper. Dealing with raw data, the application of the seasonal difference operator tends to remove more than due and could induce overdifferencing. Nevertheless, can be assumed that HEGY (Hylleberg, Hendry, Granger, Yoo, 1992) methodology leads to weak (but relevant) indication on non-stationarity. For this reason, the application of the seasonal filter could be a reasonable general practice in controlling for seasonal fluctuations. It tends to assume a prevalent deterministic character for industrial production and railways transported goods. Such evidence is confirmed also for TTP, but less significantly. Exports and imports, on the contrary, present seasonal fluctuations which strongly evolves over time (the twelve monthly dummies have been estimated to be not significant along the sample period). In that latter case, the application of annual differences could not be sufficient and residual seasonality could produce bias in estimation results. In these cases, such noise could be controlled augmenting the number of lags for dependent variable and also controlling for the trading day effect.

FREQ.	MQW(5)		IPI(3)		XQW(3)		TTP(0)		TK(4)	
0	–3.7		–3.4		–2.4		–5.3	**	–4.3	**
$\pi/6$	11.0	**	4.8		3.3	*	18.7	**	2.2	
$\pi/3$	6.1		2.6		0.6		19.2	**	7.5	**
$\pi/2$	8.4	**	7.0	*	0.1		15.6	**	12.0	**
$2\pi/3$	3.4		8.6	**	0.8		16.3	**	8.4	
$5\pi/6$	2.9		3.8	*	0.3		11.2	**	4.7	
π	–1.8	*	–1.7		–0.5		–2.9	**	–1.9	*

Table 1. Unit roots test. t-test for the 0 and π frequencies, F-test for the others. '*' and '**' indicates significance at 5% and 1%, respectively. Number of lags in parenthesis.

2.3. Cyclical components

The predictive ability of forecasting models for imports, deeply depends on the leading properties of the corresponding indicators. In this section we pay particular attention to the cyclical characteristics of the time series considered. Our aim is to observe stylized facts which could help us in the selection of the variables for model identification. As it is known, the series of import is characterized by a large and deep cyclical component, with high variability. Moreover, it shows a systematic lead on aggregate business cycle. This feature makes more difficult the selection of variables which present a regular lead on the import cycle. Cycle components of each series have been extracted applying Band-Pass filter in the form developed by Baxter and King (1999). Cyclical frequencies ranges from 18 to 96 months (up to 8 years). Extracted cycles have been plotted in Figure 1. Industrial production cycle appears to be coincident with that of imports. Correlation measures, reported in table 2, show a little lead of the latter series, according to consolidated empirical evidence. With reference to the other indicators of industrial activity, the series of railway transport of goods seems to be coincident. As it results from the figure, it shows a more defined lead of imports only in the second part of the considered time span. Production prospects represents the series showing the greater lead (two months on average) and the larger variability. The same indicator confirms previous findings about the leading properties of TTP on industrial production (about 5 months). Finally, exports series results to be coincident with imports, partly contrary to consolidated evidence.

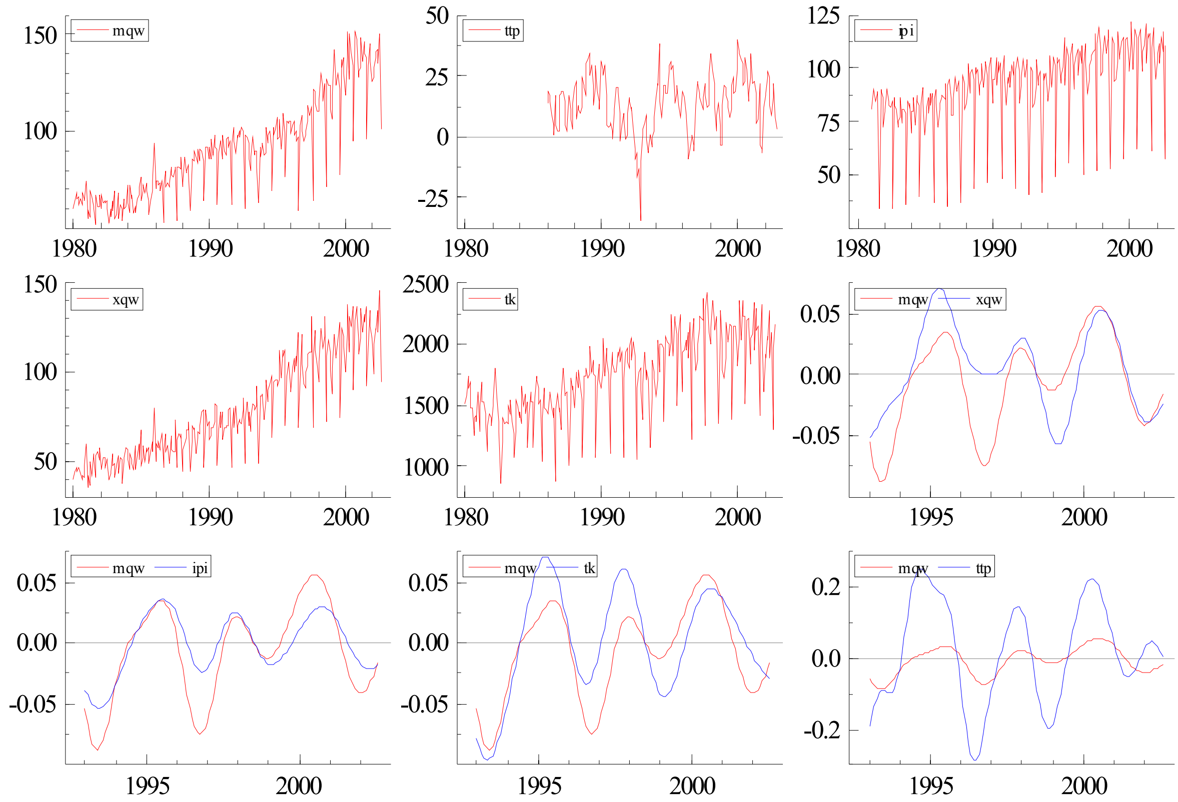

Figure 1. First five plots represent seasonally unadjusted series. In the remaining, indicators'cyclical components are plotted against that of MQW.

SERIES	σ	$\rho(0)$	$\rho(\max)$	lead(+)/lag(-)
XQW	0.034	0.678	0.678	0
IPI	0.022	0.871	0.888	−1
TK	0.041	0.794	0.794	0
TTP	0.140	0.723	0.797	+2
TTPvsIPI		0.586	0.764	+4

Table 2. Cyclical analysis. $\rho(0)$ is the correlation between the series and MQW; $\rho(\max)$ indicates the maximum cross-correlation; lead(+)/lag(-) is the interval in months at which $\rho(\max)$ is observed.

2.4. Testing for nonlinearity

Over the last period, the interest which the literature has shown for non-linear time series models has been steadily increasing. The main idea under this approach is the fact that some economic time series show the characteristic to have a non-linear mean over the period of observation. In this section, we test the null hypothesis of linearity against a well specified non-linear alternative, consisting in a particular class of *regime-switching* models, known as smooth transition autoregressive (STAR, Granger and Terasvirta, 1993).

The hypothesis testing in the STAR framework involves tests of linearity against the alternatives of LSTAR or ESTAR nonlinearity and heteroskedasticity. First of all, the testing problem is complicated by the existence of so-called nuisance parameters under the null hypothesis. The presence of such parameters causes the lack of availability of the standard statistical theory for test statistics. In particular, STAR models present parameters which are not restricted under the null. A large literature has been developed to face the problem of identifiability of nuisance parameters under the null hypothesis. In this paper we adopt the approach developed by Luukkonen, Saikkonen and Terasvirta (1988), according to which the transition function can be approximated by a Taylor series. This fact is a solution to the identification problem and linearity can be tested using usual methods and distributions. Testing is performed taking into account two different specifications of such function. The transition function is, firstly, assumed to follow a (first-order) logistic function, leading to the logistic STAR model (LSTAR); secondly, the exponential function, getting the exponential STAR (ESTAR). In the case the alternative is assumed to be LSTAR nonlinearity, the auxiliary regression becomes

$$y_t = \beta_0' x_t + \beta_1' x_t s_t + \varepsilon_t \qquad (1)$$

Testing for linearity is equivalent to test the null hypothesis $H_0'': \beta_1 = 0$. This test statistics can be constructed as an LM type test with a χ^2 distribution with *p + 1* degrees of freedom under the null hypotesis of linearity. It is reported as LM1 statistics in the following tables. A variant of this statistics has been developed by Luukkonen *et al.* (1988), since he noticed LM1 has no power when only the intercept differs across regimes. A third order approximation of the logistic transition function yields the auxiliary regression

$$y_t = \beta_0' x_t + \beta_1' x_t s_t + \beta_2' x_t s_t^2 + \beta_3' x_t s_t^3 + \varepsilon_t \qquad (2)$$

In such equation, the test for null hypothesis, named as LM3, reduces to $H_0'': \beta_1 = \beta_2 = \beta_3 = 0$, which again can be tested by a standard LM type test. Testing against an ESTAR alternative, assuming the transition function as an exponential, requires an auxiliary equation constructed on the basis of a second-order Taylor approximation of the type

$$y_t = \beta_0^{'} x_t + \beta_1^{'} x_t s_t + \beta_2^{'} s_t^2 + \beta_3^{'} s_t^3 + \beta_4^{'} s_t^4 + \varepsilon_t \tag{3}$$

The null hypothesis to be tested is $H_0^{'}: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$. To evaluate empirical findings we use the F-version of the LM test statistics, denoted LM4 in the following tables. Especially in small samples, it is a good strategy to use F-version of LM-type tests as it results much more robust than χ^2 variant. Both tests can be constructed making use of two auxiliary regressions. For LM4, the F-version is

$$LM_4 = T(SSR_0 - SSR_1)/SSR_0 \tag{4}$$

where SSR_1 and SSR_0 represent, respectively, the residual sum of squares of the OLS estimates with and without interaction terms.

All the variables are taken in seasonal differences; the transition variables we consider are lagged values of such series. The maximum value of the delay parameter is equal to 6. Finally, deterministic trend has been considered as an additional transition variable. Auxiliary equations have been specified with constant, trend, eleven seasonal dummies and the level of the endogenous variables lagged up to the 12^{th} lag. First, concentrating on standard evidence (LM1 test), it emerges significant nonlinearity for all variables considered, with transition function lagged by 1 to 3. Import and production prospects make exception. Under a more robust evidence (LM3 test), only nonlinearity at the 10 % significance level is detected for import and industrial production.[3] Weak evidence of ESTAR nonlinearity has been found for export and TK. No evidence appears in the case of linear and deterministic transition function. Considering heteroskedasticity robust tests, all the variables satisfy the linearity assumption, for whatever specification of the transition function. As a preliminary conclusion, any non linearity present in the series may be considered as moderate since the linearity tests does not reject the null with strong evidence.

[3]We are considering transition variable lagged by 2 and by 3 and specified as logistic.

	Standard test			Heteroskedasticity robust test		
	LM1	LM3	LM4	LM1	LM3	LM4
trend	0.966	1.000	1.000	0.848	1.000	0.998
d12y(-1)	0.706	0.260	0.508	0.937	0.663	0.859
d12y(-2)	0.805	0.052	0.125	0.935	0.260	0.869
d12y(-3)	0.457	0.030	0.127	0.685	0.530	0.807

Table 3. LM-type test for STAR non-linearity: IMPORT OF GOODS

	Standard test			Heteroskedasticity robust test		
	LM1	LM3	LM4	LM1	LM3	LM4
Trend	0.642	0.255	0.541	0.783	0.965	0.984
d12y(-1)	0.038	0.150	0.013	0.296	0.644	0.879
d12y(-2)	0.046	0.114	0.350	0.432	0.845	0.968
d12y(-3)	0.107	0.140	0.181	0.629	0.959	0.993

Table 4. LM-type test for STAR non-linearity: EXPORT OF GOODS

	Standard test			Heteroskedasticity robust test		
	LM1	LM3	LM4	LM1	LM3	LM4
Trend	0.582	0.952	0.879	0.725	0.962	0.897
d12y(-1)	0.011	0.248	0.070	0.238	0.796	0.823
d12y(-2)	0.032	0.504	0.744	0.390	0.949	0.994
d12y(-3)	0.053	0.518	0.359	0.436	0.917	0.994

Table 5. LM-type test for STAR non-linearity: INDUSTRIAL PRODUCTION

	Standard test			Heteroskedasticity robust test		
	LM1	LM3	LM4	LM1	LM3	LM4
Trend	0.583	0.998	0.974	0.422	0.973	0.995
d12y(-1)	0.035	0.126	0.225	0.056	0.695	0.833
d12y(-2)	0.025	0.130	0.419	0.078	0.992	0.958
d12y(-3)	0.018	0.064	0.224	0.107	0.968	0.778

Table 6. LM-type test for STAR non-linearity: TONS/Km OF GOODS TRANSPORTED BY RAIL

	Standard test			Heteroskedasticity robust test		
	LM1	LM3	LM4	LM1	LM3	LM4
Trend	0.977	1.000	1.000	0.724	0.991	0.990
d12y(-1)	0.434	0.655	0.732	0.442	0.872	0.813
d12y(-2)	0.022	0.315	0.058	0.169	0.840	0.946
d12y(-3)	0.119	0.117	0.091	0.274	0.956	0.958

Table 7. LM-type test for STAR non-linearity: PRODUCTION PROSPECTS

3. THE FORECASTING MODEL

An explicit aim of this work is to find out a reliable and simple model to forecast the Italian import of goods, in order to overcome the problems arising from the delay with which official information is developed. Empirical evidence on the non-linearity of the series suggests to consider linear models.

This methodology, with simple seasonal component, offers advantages over more complicated ones in terms of their short-term forecasting accuracy. Nevertheless, the single equation framework offers an oversimplified option and does not allow for multi-step dynamic forecasts. For all these reasons we decide to consider the well established VAR (Vector Autoregression) framework.

As we shown in Section 2, the four time series that we consider have different seasonal properties, but there was a strong evidence that the presence of a complete set of unit roots has to be rejected. This implies that if we parameterize the VAR in seasonal differences, we are likely to over-difference the series. Nevertheless, there is some evidence on the effect on forecasting performance deriving from imposing all the seasonal roots at unity when this is not the case in reality. There are indications that filtering out only the correct unit roots, in general, does not produce superior forecasts. In particular, Lyhagen and Löf (2001) suggest that when the model is not known and the aim of the modelling exercise is forecasting, a VAR in annual differences may be a better choice than a seasonal error correction model based on seasonal unit roots pre-testing. Moreover, Osborn, Heravi and Birchenhall (1999) find that, despite the series typically providing evidence against seasonal integration, models based on seasonal differences produce forecasts that are at least as accurate as those based on deterministic seasonality. Clemens and Hendry (1997) conclude, from their analysis, that imposing seasonal unit roots and using the model based on seasonal differences may improve accuracy even if the imposition is not warranted according to the outcomes of unit root tests. As it has been shown also in Paap et al. (1997), models based on seasonal differences improve forecast even in presence of structural breaks occurring during the forecast period. Therefore, we parameterize our VAR in seasonal differences.

The model has been specified with reference to the non seasonal adjusted series and, in its more general formulation, takes the form:

$$\Delta\Delta_{12} y_t = \beta' \Delta_{12} y_{t-1} + \sum_{j=1}^{13} \gamma_j' \Delta\Delta_{12} y_{t-j} + \phi' d_t + \varepsilon_t \qquad (5)$$

where $\Delta = (1-L)$, $\Delta_{12} = (1-L^{12})$, L is the usual lag operator such that $L^p_{z_t} = z_{t-p}, y_t = (IPI_t, TK_t, TTP_t, XQW_t)'$, and d_t are the deterministic components.

As discussed in the previous sections, we have considered four variables for the forecasting exercise. These variables are those which best represent the path of the series we want to forecast. Combining in different ways these variables we have obtained four VAR models which we, now, describe in detail.

The first one of them (VAR1) consists of three variables: imports, which is the series that we are trying to forecast, exports and the series of goods transported on railways. The deterministic part considers the correction for working days and two seasonal dummies (January 1993 and December 1996).

The second one (VAR2) is made up of imports, the series of tons of goods transported on railways, and the series of future production prospect released monthly by ISAE. While the first two equations of the VAR model include the trading days, the third one contains only a seasonal dummy (November 1992).

The third model (VAR3)is formed by the series of imports, exports and the industrial production index released monthly by ISTAT. Trading days are contained in the deterministic part. It has not been necessary to use any dummy variable.

The last one (VAR4) is different from the others. It is made up of four variables: imports, exports, tons of goods transported on railways and future production prospect. It contains, in its deterministic part, the working days' correction.[4]

The large use of dummy we have made has its practical justification in the presence of many outliers in the data. Moreover the use of dummies to correct the anomalous data is a very spread practice in this kind of literature. The deterministic part include also, in some cases, the variables $\Delta_{12} log(TD_t)$ and $\Delta_{12} log(TD_{t-1})$, with TD_t the number of trading days in month t. The number of trading days significantly influences manufacturing activity. The use of the lagged value is not very common in practice, but, in presence of particularly

[4]We have used the expressions working days' and trading days' corrections with the same significance only for simplicity.

unfavorable (favorable) trading days configurations, it is legitimate to expect that firms tend to compensate lower (higher) realized production in the following month. In order to include the working days' correction in the deterministic part of the models, we have tested the significance of such regressors in each equation of the VAR models. The estimated coefficients of the two variables, are, generally, both highly significant, and seem to confirm this point of view.

As it is known, models like these are subject to the course of dimensionality: the number of parameters grows as the square of the number of variables times the maximum lag contained in the more general specification of the model. For this reason the VAR models have been sequentially simplified to obtain a more parsimonious parameterization using a "general-to-specific" reduction (Krolzig, 2000; Krolzig and Hendry, 2001). Nevertheless, by reducing the complexity of the VAR, it is necessary to ensure, simultaneously, that the parsimonious subset VAR will contain all the information embodied in the unrestricted one. To assure this, in each step of the reduction, a statistical test is made. The reduction procedure stop when it is not possible to eliminate another variable without loosing information from the general model. Table 8 reports the lags' structure after the reduction.

	VAR1	**VAR2**	**VAR3**	**VAR4**
Lags' structure	1,3,4,6,12,13	1,2,3,5,8,12	1,2,5,9	1,2,3,5,12

Table 8. Lags' structure after the reduction

Only one model represents a significative parsimonious version of the more general one. In many cases, even if the restricted VAR is more parsimonious that the starting one, it is still rather highly parameterized including a large number of lags. It is interesting to notice that in all, or nearly, the restricted specifications are present the lags from 1 to 3, which capture the autoregressive components of the model, and the lag 12, which is characterized from the seasonal components.

3.1. Reduction diagnostics

The main statistics and diagnostics of the VAR models estimated over the period 1990.01-1999.05 are reported in the following tables. The tables report the standard error of each equation in the VAR (σ), the correlation of actual and fitted values (ρ), the p-value of the LM test for residuals autocorrelation up to the twelfth order (AR 1-12), and the p-value of the test for residual normality (Normality). The lower part of the tables reports the p-values, in their F-form, of the parameter constancy forecast tests. The first one of them does not consider parameter uncertainty.

	σ	ρ	***AR* 1—12**	***Normality***
dd12lmqw	0.040	0.867	0.737	0.163
dd12lxqw	0.048	0.816	0.167	0.253
dd12tk	0.101	0.718	0.776	0.428
VAR				0.355
Parameter stability test on the forecasting interval 1997.06 – 1999.05				
F_{Ω}	0.287			
$F_{V_{(e)}}$	0.639			
$F_{V_{(E)}}$	0.766			

Table 9. Main VAR diagnostic: estimation period 1990.01 – 1999.05

	σ	ρ	***AR* 1—12**	***Normality***
dd12lmqw	0.042	0.860	0.864	0.666
dd12lxqw	0.048	0.849	0.545	0.105
dd12tk	0.098	0.760	0.568	0.281
VAR				0.314
Parameter stability test on the forecasting interval 1997.06 – 1999.05				
F_{Ω}	0.225			
$F_{V_{(e)}}$	0.621			
$F_{V_{(E)}}$	0.658			

Table 10. Main VAR diagnostic: estimation period 1990.01 – 1999.05

All the VAR's reduced specifications show good diagnostics over the considered estimation period. In particular, the correlation of actual and fitted values seems to be a strong result. In general, this diagnostic satisfies our expectation about the good specification of the models. The tests for parameter consistency, calculated over the forecast evaluation sample, do not reject structural stability. They seem to be very robust except for the first one (F_{Ω}, which does not consider parameter uncertainty).

	σ	ρ	***AR* 1—12**	***Normality***
dd12 lmqw	0.056	0.723	0.284	0.145
dd12 lxqw	0.051	0.759	0.159	0.173
dd12 tk	0.035	0.827	0.186	0.758
VAR				0.383
Parameter stability test on the forecasting interval 1997.06 − 1999.05				
F_{Ω}	0.992			
$F_{V_{(e)}}$	0.998			
$F_{V_{(E)}}$	0.999			

Table 11. Main VAR diagnostic: estimation period 1990.01 − 1999.05

	σ	ρ	***AR* 1—12**	***Normality***
dd12 lmqw	0.045	0.852	0.240	0.736
dd12 lxqw	0.043	0.856	0.125	0.414
dd12 tk	0.100	0.779	0.895	0.176
dd12ttp	0.044	0.888	0.826	0.665
VAR				0.675
Parameter stability test on the forecasting interval 1997.06 − 1999.05				
F_{Ω}	0.183			
$F_{V_{(e)}}$	0.722			
$F_{V_{(E)}}$	0.798			

Table 12. Main VAR diagnostic: estimation period 1990.01 − 1999.05

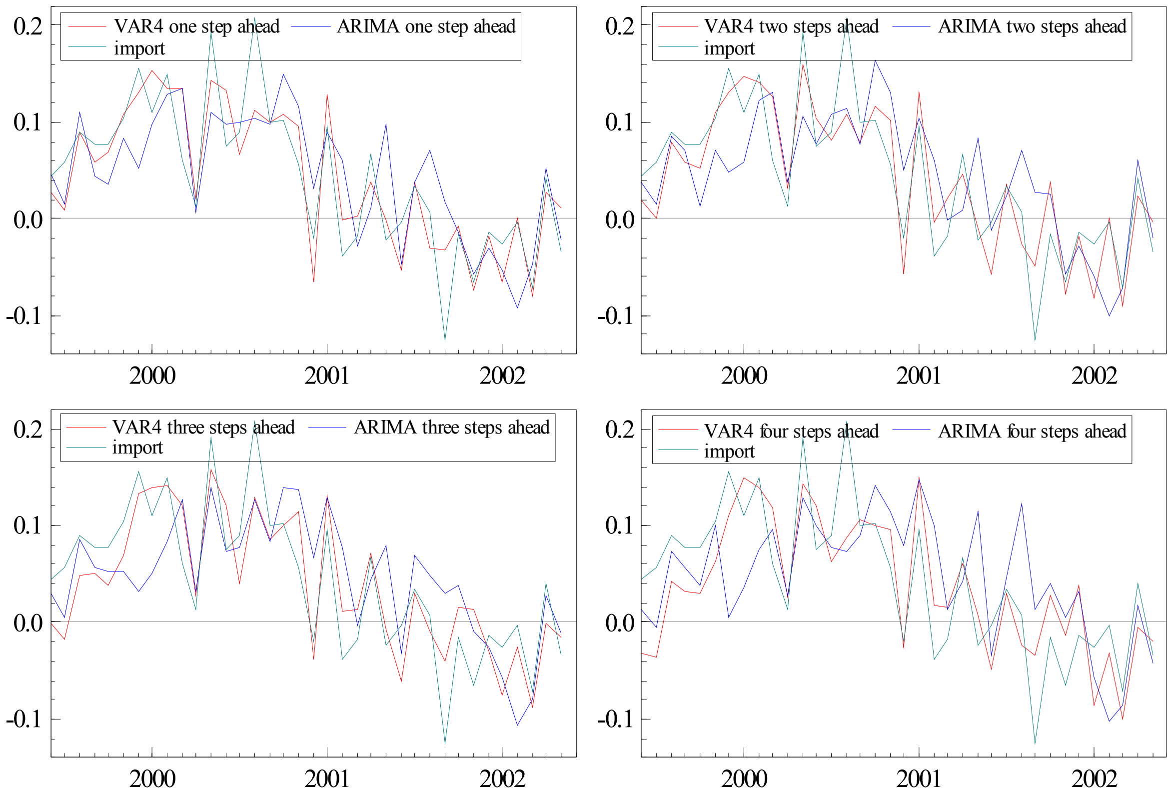

Figure 2. Forecast: graphical analysis.

4. FORECAST EVALUATION

In this section we evaluate the forecasting ability of the four VAR models as opposed to an ARIMA model. We are interested to investigate not only if all the VAR models offers a best prediction with respect to the benchmark produced by the ARIMA, but also which of the VAR models produce a best prediction of the Italian import of goods. To make the evaluation more interesting, the ARIMA model has been enriched with a deterministic part that includes trading days and Easter effects. This model has been estimated recursively by maximum likelihood and the forecasts have been produced using TRAMO:[5] a model like this, represent a good benchmark very difficult to exceed. The comparison has been made over a fairly long period (1999:7-2002:5).

[5]For more details see Gómez and Maravall (1998), Maravall (1995).

STEPS AHEAD	1	2	3	4
VAR1	4.33	4.71	5.26	6.05
	3.51	3.79	4.40	5.29
VAR2	3.83	3.95	4.37	4.92
	2.82	3.15	3.65	4.08
VAR3	4.03	4.37	4.80	5.31
	3.14	3.28	3.86	4.32
VAR4	3.75	3.68	4.13	4.83
	2.87	2.99	3.49	4.10
ARIMA	5.57	5.85	6.00	7.01
	4.14	4.47	4.74	5.52

Table 13. Forecast evaluation diagnostics. For each model the first row reports the root mean square error (RMSE), the second one indicates the mean absolute error (MAE)

In Table 13, the root mean square error and the mean absolute error in each steps ahead of the forecast are reported for each VAR model and for the benchmark produced with the ARIMA model. It is easy to see that each VAR model presents best results in term of prediction evaluation if compared with the ARIMA. In some cases, the difference began very large (in the order of two points percent) in each step. The best performance seems to be that of the VAR model made up of four variable (VAR4). At the first step, the root mean square error of this VAR model (3.75 percent) is about two point per cent above of the ARIMA's one. This difference remain in all the other steps. Similar consideration could be made about the mean absolute error.

5. FORECAST ENCOMPASSING TESTS

In order to compare the predictive accuracy of our models, we have performed some econometric tests. The forecast encompassing tests are a way to compare the quantity of information contained in two models. In other terms, if a model encompasses another, we can not conclude that it has a great quantity of information, but only that it contains a part of information that is not in the other model. Nevertheless, it could be true that the second model (the one which is encompassed), contains itself a quantity of information which is not contained in the first (the one which encompasses). For this reason, it is necessary to compare the models one with another and viceversa. In this paper, we have used two different ways to perform forecast encompassing tests: one was firstly proposed by Diebold and Mariano (1995), the second, is due to Fair and Shiller (1990). In the following, we shortly explain the way in which the two different tests perform.

Diebold and Mariano

Suppose one has two series of n forecasts each to be compared. Let $\{e_{it}\}_{t=1}^{n}$ be $h-$step ahead forecast error deriving from model i. Denote by $d_t = e_{it}(e_{it} - e_{jt})$ an arbitrary function. The null hypothesis of equality of expected forecast performance is $E(d_t) = 0$. It is natural to consider $\bar{d} = n^{-1}\sum_{t=1}^{n} d_t$, so that $\sqrt{n}(\bar{d} - \mu_d) \xrightarrow{d} N(0, 2\pi f_d(0))$, where μ_d is the population mean of d_t, and $f_d(0)$ is the spectral density of d_t at frequency zero. Diebold and Mariano (1995) propose basing the test of equal forecasting accuracy on

$$DM = \frac{\bar{d}}{\sqrt{n^{-1}2\pi\widehat{f_d(0)}}} \tag{6}$$

which, under the null, tends to a standardized normal distribution when $\widehat{f_d(0)}$ is a consistent estimate of $f_d(0)$. In order to correct for the size distortions noticed in the test based on DM, Harvey et al. (1989,1997,1998) propose modifying the test in this way:

$$DM^* = \left(\frac{n+1-2h+n^{-1}h(h-1)}{n}\right)^{\frac{1}{2}} DM. \tag{7}$$

Using this statistics comport that, under the null, forecast i encompasses forecast j and $E(d_t) = 0$. On the other hand, under the alternatives, forecast i could be improved by incorporating some of the features present in forecast j. In this paper, we have used the DM^* version of the test. In order to obtain a consistent estimate of $f_d(0)$, we use an unweighted sum of the sample autocovariances up to $h-1$, of the form:

$$2\pi\widehat{f_d(0)} = \hat{\gamma}_0 + 2\sum_{\tau=1}^{h-1}\hat{\gamma}_\tau,$$

with γ_k the lag$-k$ sample autocovariance.

Fair and Shiller

To better compare the performance of the models, we also propose another exercise of forecast encompassing. This second approach we refer to, was firstly proposed by Fair and Shiller (1990). The main idea is to compare the information contained in the forecast produced by the different models taken two at a time. In other words, the approach is based on the comparison of the series forecasted across different models only through a simple regression of the form:

$$\Delta_{12} y_t = \alpha + \beta_{M1}(\hat{y}_{M1,t} - y_{t-12}) + \beta_{M2}(\hat{y}_{M2,t} - y_{t-12}) + e_t, \qquad (8)$$

where $\hat{y}_{M1,t}$ and $\hat{y}_{M2,t}$ are the forecast obtained using model $M1$ and $M2$ respectively. If only one of the two models contains relevant information, the corresponding estimated coefficient will be significant.

M_i/M_j	**ARIMA**	**VAR1**	**VAR2**	**VAR3**	**VAR4**
ARIMA		2.551	3.244	3.406	3.544
		(0.015)	(0.002)	(0.001)	(0.001)
VAR1	2.527		3.047	2.910	3.364
	(0.016)		(0.004)	(0.006)	(0.001)
VAR2	0.121	3.756		3.190	0.435
	(0.904)	(0.000)		(0.002)	(0.665)
VAR3	-0.596	0.559	1.208		1.473
	(0.554)	(0.579)	(0.234)		(0.149)
VAR4	0.220	0.471	0.095	0.956	
	(0.826)	(0.640)	(0.924)	(0.345)	

Table 14. Modified Diebold-Mariano test. DM^* statistics evaluated for the one-step ahead forecast and their p-value under the null are reported. The null hypothesis is that the forecasts produced by model M_i encompass those produced by model M_j.

Main results of forecast encompassing tests

In this section we look at the main results of the forecast encompassing tests (Tables 14 - 17).

The test of forecast encompassing based on Diebold and Mariano statistics, made up only for one and two steps ahead, show very strong results. The null hypothesis is that the forecasts produced by model M_i encompass those produced by model M_j. From our viewpoint it seems relevant to note that in no cases the ARIMA projections seem to embody some piece of information that would be useful for improving both the one-step and the two-step ahead VAR forecast. Conversely, except for the VAR1, our models contains information not contained in the ARIMA forecast. In some cases, as for the

VAR4, this evidence is particulary strong, but, in general, seems to be strongly confirmed by the p-values. Between our models, the VAR4, is still that which contains the greater part of information; in fact, it encompasses all the other models, while is encompassed only by the VAR2. Thus, the latter result has to be considered very anomalous as the VAR2 encompasses only the VAR4 and not encompasses the other VAR models at one-step ahead forecast. At the shorter horizon (one step ahead), the VAR2 seems to be the once which incorporate less information with respect to the others. At the longer horizon (two steps), it is more difficult to do similar consideration.

M_i/M_j	**ARIMA**	**VAR1**	**VAR2**	**VAR3**	**VAR4**
ARIMA		3.214	4.051	4.108	4.806
		(0.002)	(0.000)	(0.000)	(0.000)
VAR1	0.496		3.649	3.090	5.683
	(0.622)		(0.000)	(0.003)	(0.000)
VAR2	0.718	1.630		2.339	0.947
	(0.477)	(0.111)		(0.025)	(0.349)
VAR3	-0.611	0.366	1.804		1.985
	(0.545)	(0.716)	(0.079)		(0.054)
VAR4	0.225	0.468	-0.453	0.623	
	(0.822)	(0.642)	(0.652)	(0.537)	

Table 15. Modified Diebold-Mariano test. DM^* statistics evaluated for the two-steps ahead forecast and their p-value under the null are reported. The null hypothesis is that the forecast produced by model M_i encompass those produced by model M_j.

M_1/M_2		**ARIMA**	**VAR3**	**VAR2**	**VAR1**
	α	0.10			
VAR3	β_{M1}	2.42			
	β_{M2}	−1.85			
	α	−0.22	0.27		
VAR2	β_{M1}	2.40	−0.08		
	β_{M2}	−1.89	0.47		
	α	0.20	0.33	0.45	
VAR1	β_{M1}	2.48	1.12	1.29	
	β_{M2}	−1.67	−0.61	−0.75	
	α	0.10	0.20	−0.12	0.28
VAR4	β_{M1}	1.47	−0.93	−1.01	−1.57
	β_{M2}	−0.98	1.37	1.34	2.12

Table 16. Predictive accuracy tests (Fair-Shiller): one step ahead forecast

M_1/M_2		ARIMA	VAR3	VAR2	VAR1
	α	0.35			
VAR3	β_{M1}	2.38			
	β_{M2}	−1.51			
	α	0.04	0.51		
VAR2	β_{M1}	2.05	−0.44		
	β_{M2}	−1.33	1.01		
	α	0.53	0.62	0.74	
VAR1	β_{M1}	2.66	1.38	1.81	
	β_{M2}	−1.42	−0.66	−0.95	
	α	0.21	0.21	−0.56	0.50
VAR4	β_{M1}	0.48	−2.41	−2.62	−2.75
	β_{M2}	0.15	3.05	3.06	3.63

Table 17. Predictive accuracy tests (Fair-Shiller): two step ahead forecast

The results for the two-step ahead are very similar and could be commented in the same way. This evidence is only in part confirmed by the approach due to Fair and Shiller. In fact, using this methodology, the VAR4 does not encompass the ARIMA benchmark, which is, otherwise, encompassed by all the other VAR models.

CONCLUDING REMARKS

In this paper, we evaluate several econometric models performing short-horizon forecasts of Italian imports of goods. Year-to-year growth rate of the monthly seasonally unadjusted series is the variable to predict. For the Italian case, imports mainly consist of raw materials and intermediate goods. For this reason, it can be taken as a significant leading indicator of the aggregate business cycle. This feature, though extremely useful in assessing very short-run dynamics of Italian economy, cannot be properly exploited due to the lack of the availability of the statistical information.

Preliminary analysis on the series used in our forecast exercise have been carried out. Among them, a set of nonlinearity tests have been applied to the series of Italian total imports taken in seasonal differences. The null, that the process is a linear autoregression, has been tested against the alternative of a nonlinear self-exciting threshold autoregressive model (SETAR). A strong evidence of linearity of the series of import has appeared, so stylized VAR (VectorAutoregression) models have been specified along a restricted set of variables (industrial production, exports among others). VAR forecasting ability has been evaluated as opposed to that of a linear univariate benchmark (ARIMA) model. Main forecast diagnostics and two types of forecast encompassing tests have been presented. The principal result obtained concerns the superiority of the VAR models which systematically outperform the ARIMA benchmark model.

REFERENCES

Baxter M. and R.G. King (1999), "Measuring Business Cycles: Approximate Bandpass Filter for Economic Time Series", *Review of Economics and Statistics*, vol. 81, pp. 575-593.

Bealieu J.J. and J.A. Miron, (1993), "Seasonal Unit Roots in Aggregate U.S. Data", *Journal of Econometrics*, vol. 55, pp. 305-328.

Bruno G. and C. Lupi (2001), "Forecasting Iindustrial Production and the Early Detection of Turning Points", *Working Papers, ISAE* n. 20/01.

Clements M.P. and D.F. Hendry (1997), "An Empirical Study of Seasonal Unit Roots in Forecasting", *International Journal of Forecasting*, vol. 13, pp. 341-355.

Diebold F.X. and R.S. Mariano (1995), "Comparing Predictive Accuracy", *Journal of Business and Economic Statistics*, vol. 13, pp. 253-263.

Fair C.R. and R.J. Shiller (1990), "Comparing Information in Forecasts from Econometric Models", *American Economic Review*, vol. 80 , pp. 375-389.

Gómez V. and A. Maravall, (1998), *Programs TRAMO and SEATS: Instructions for the User*, Madrid:Banco de España.

Granger C.W.J. and T. Terasvirta (1993), *Modelling Nonlinear Economic Relationships,* Oxford: Oxford University Press.

Harvey A.C. (1989), *Forecasting Structural Time Series Models and the Kalman Filter,* Cambridge: Cambridge University Press.

Harvey A.C., S.J. Leybourne and P. Newbold (1997), "Testing the Equality of Prediction Mean Squared Errors", *International Journal of Forecasting*, vol. 13, pp. 281-291.

Harvey A.C., S.J. Leybourne and P. Newbold (1998), "Test for Forecast Encompassing", *Journal of Business and Economic Statistics*, vol. 16, pp. 254-259.

Hylleberg S., R.F Engle, C.W.J. Granger and B.S. Yoo (1990), "Seasonal Integration and Cointegration", *Journal of Econometrics*, vol. 44, pp. 215-238.

Krolzig H.M. (2000), "General-to-Specific Reductions in Vector Autoregressive Processes", *Economics Discussion Paper 2000-W34 Nuffield College Oxford.*

Krolzig H.M. and D.F. Hendry (2001), "Computer Automation of General-to-Specific Model Selection Procedures", *Journal of Economic Dynamics and Control,* pp. 831-866.

Luukonen R., P. Saikkonen and T. Terasvirta (1988), "Testing Linearity Against Smooth Transition Autoregressive Models", *Biometrika,* vol. 75, pp. 491-499.

Lyhagen J. and M. Löf (2001), "On Seasonal Error Correction When the Process Include Different Numbers of Unit Roots", *SSE/EFI Working Paper Series in Economics and Finance,* No. 0418.

Maravall A. (1995), "Unobserved Component in Economic Time Series", In H. Pesaran et al. (Ed.), *The Handbook of Applied Econometrics*, vol. 1. *Oxford: Basil Blackwell*, pp. 12-72.

Osborn D.R. , S. Heravi and C.R. Birchenhall (1999), "Seasonal Unit Roots and Forecasts of Two-Digit European Industrial Production", *International Journal of Forecasting*, vol. 15, pp. 27-47.

Paap R., P.H. Franses and H. Hoek (1997), "Baesyan Analysis of Seasonal Unit Roots and Seasonal Mean Shifts", *Journal of Econometrics*, vol. 78, pp. 359-380.