ISAE Istituto di studi e analisi economica

FORECASTING INDUSTRIAL PRODUCTION AND THE EARLY DETECTION OF TURNING POINTS

by

Giancarlo Bruno and Claudio Lupi

ISAE

Rome

June 2001

ABSTRACT

In this paper we propose a simple model to forecast industrial production in Italy. We show that the forecasts produced using the model outperform some popular forecasts as well as those stemming from a trading days- and outlier-robust ARIMA model used as a benchmark. We show that the use of appropriately selected leading variables allows to produce up to twelve-step ahead reliable forecasts. We show how and why the use of these forecasts can improve the estimation of a cyclical indicator and the early detection of turning points for the manufacturing sector. This is of paramount importance for short-term economic analysis.

JEL Classification: C53, C32, E32.

Key words: Forecasting, VAR Models, Industrial Production, Cyclical Indicators.

The authors would like to thank Sergio de Nardis for comments and discussions. The usual disclaimer applies.

NON TECHNICAL SUMMARY

This paper deals with the issue of forecasting the Italian industrial production index. While there are many of such predictions produced by several research institutes, either they are concerned with a very short horizon (up to two months ahead) or they produce just annual figures. Here the authors propose using some leading variables to build a multivariate model so as to allow reliable monthly forecasts to be produced up to twelve months ahead.

Multi-step forecasting of industrial production index is useful with respect to many aspects. In fact, despite the growing importance of the service sector, industrial production is still important in explaining aggregate business cycle fluctuations. Forecasts of industrial production can then be used as an input in larger models, which are often criticized for their (in)ability in tracking business cycles turning points.

Furthermore, the authors show that the use of the multi-step forecasts reduces dramatically the uncertainty in estimating a cyclical indicator from the industrial production index at the very end of the series, which is a fundamental issue in short-term economic analysis.

PREVISIONE DELL'INDICE DELLA PRODUZIONE INDUSTRIALE E INDIVIDUAZIONE DEI PUNTI DI SVOLTA

SINTESI

In questo lavoro viene proposto un modello previsivo dell'indice della produzione industriale per l'Italia. Le previsioni prodotte da tale modello si dimostrano superiori a quelle derivanti da un modello ARIMA, corretto per tenere conto di valori anomali ed effetti di calendario, utilizzato come *benchmark*. L'uso di opportune variabili anticipatrici consente di ottenere previsioni attendibili per un orizzonte previsivo di dodici mesi. Inoltre, si mostra come l'utilizzazione di tali previsioni migliori sensibilmente la stima di un indicatore ciclico della produzione industriale e anticipi notevolmente l'individuazione dei punti di svolta, fattori, questi, di primaria importanza per l'analisi economica di breve termine.

Classificazione JEL: C53, C32, E32.

Parole chiave: Previsioni, modelli vettoriali autoregressivi, produzione industriale, indicatori ciclici.

INDEX

"...Economists have forecasted 9 out of the last 5 recessions." (Anonymous)

1. INTRODUCTION

Forecasting the industrial production index is an important task in short-term economic analysis. This is still true in the nowadays economies where services are undertaking an increasing weight. In fact, the industrial sector is still important in explaining aggregate fluctuations, also because some of the services activities (business services) are closely linked to the industrial ones. In addition, forecasts of industrial production can be useful in more general forecasting models. Bovi *et al*. (2000), for example, use the industrial production index in a closing equation of a demand-based forecasting model for quarterly GDP. In this case, reliable three-month ahead forecasts would be extremely useful. Furthermore, cyclical indicators of the manufacturing sectors may be derived from the industrial production index series: this is commonly done by applying signal extraction techniques, and accurate forecasts of the series to be filtered are essential.

From a general standpoint it should be said that most of the existing models in Italy offer an early estimation of the industrial production, rather than a pure forecast. In fact, the official indicator is released by the National Statistical Institute (ISTAT) 45 days after the end of the reference month, so that a two-step ahead prediction is necessary to achieve a *nowcast* of the indicator itself. This is the case, in particular, for two popular predictions released monthly by CSC and IRS,[1] respectively. In the second half of month t (when the official indicator is available up to month $t-2$), CSC releases a preliminary survey-based estimate of month t and a revised estimate for month $t-1$. A similar dissemination scheme is followed also by IRS, which however uses a model based on electricity consumption to produce its projections; at the half of month t a preliminary estimate of the same month is released, and a final one is published at the beginning of month $t+1$.

In this paper, we propose a simple model able to produce satisfactory forecasts of the industrial production index well beyond the two-step ahead *nowcasts.* We show that the projections deriving from a simple VAR using appropriately selected leading variables can well compete with the two aforementioned accredited forecasts in terms of predictive ability. We also show how our projections can be used

[1] CSC (Centro Studi Confindustria) is the research department of Confindustria, the Confederation of Italian Industry. IRS (Istituto per la Ricerca Sociale) is an independent no-profit social research centre. Speaking of "models" when referring to CSC projections is, strictly speaking, inappropriate, given that they are derived from a survey. However, for brevity we will refer to the different forecasting devices as "models". This should cause no confusion.

successfully to reduce uncertainty in the estimation of a cyclical indicator. Finally, we use our predictions also to improve substantially on the timely detection of turning points in the manufacturing sector. We actually think that this is a major result of this paper. Though the empirical analysis is carried out on Italian data, we feel that the implications are far reaching and the arguments developed in the paper are potentially of interest to an international audience.

The paper is organized as follows: in the next section we illustrate some preliminary analyses carried out on the time series used in the forecasting exercise; Section 3 describes the forecasting model and Section 4 is devoted to the evaluation of its predictive ability, also in comparison with the competing predictions released by CSC and IRS. Section 5 analyses the improvements deriving from using our forecasts in estimating a cyclical indicator for the manufacturing sector, and in turning point detection. The final Section concludes.

2. PRELIMINARY ANALYSIS

One of the first logical steps in a modelling strategy is the review of the available information. Econometric models already available in Italy to forecast industrial production mainly use coincident indicators of industrial activity, such as electricity consumption (see *e.g.* Marchetti and Parigi, 2000), which have the advantage of an earlier release with respect to the industrial production index, making it possible to formulate up to two-period ahead predictions (nowcasts).

The goal of obtaining "true" forecasts of the industrial production index (IPI), makes it necessary to forecast the official figures at least three months ahead. This is why it might be sensible to restrict the choice of the variables which will enter the forecasting model among those characterized by a leading pattern, discarding those which are roughly coincident. A comprehensive analysis of the properties of many Italian economic time series has been carried out by Altissimo *et al.* (1999). In part using the results contained in that paper, and after restriction of a considerably higher number of candidates, we find that two variables seem particularly interesting as potential predictors of the industrial production in Italy: the ISAE business surveys series[2] of future production prospects (PP) and the quantity of

[2] See Pappalardo (1998) and the references therein for a description of the uses of ISAE business surveys in forecasting models of Italian industrial production.

goods transported by railways (*TON*).[3] The first variable represents industrial entrepreneurs' opinions about future production. More precisely, the entrepreneurs are asked if the production in the following three-four months is expected to go "up", to be "stable", or to go "down". The answers are then synthesized with a *balance*, *i.e.* the share of "up" less the share of "down" answers. The variable obtained (PP) is therefore bounded in the interval [-100, 100], and it is a natural candidate in a forecasting model given its timely availability,[4] its explicit link with the variable to be forecast, and its long lead over the industrial production series. The usefulness of the second variable in a forecasting model is due to the fact that the merchandises transported by rail are mainly intermediate goods and raw materials used as inputs by manufacturing industries. Indeed, this variable is characterized by a fairly stable lead over the industrial production index, as well as by a short delay in its availability.

In the next sub-section a preliminary analysis of the univariate characteristics of the series of interest is presented. A log transformation is used for the series *IPI* and *TON* while *PP* is rendered unbounded using the transform[5]

$$-\log\left(\frac{200}{PP+100}-1\right). \tag{1}$$

2.1 Description of the series

Figure 1 plots the series used in this paper. They are characterized by rather heterogeneous patterns. The industrial production index shows a strong seasonality, with some cyclical fluctuations around an upward trend. About the same can be said of the railways transport of goods, which seems to feature some more pronounced cyclical movements, together with some possible outliers. A rather different pattern can be seen looking at the graph of the production expectations of industrial firms: in this case the cyclical movements are clearly predominant with respect to the other components, even though some seasonality seems to be present.

[3] The industrial production index (*IPI*) is released monthly by ISTAT, the Italian National Statistical Institute. Press releases and recent data can be found at `http://www.istat.it/Anews/proind.html`. Future production prospects (forecasts, *PP*) are released monthly by ISAE, the Institute for Studies and Economic Analyses. Recent data and updates can be found at `http://www.isae.it/ english.html`. The time series for tons of goods transported by railways (*TON*) and its updates are kindly provided by Ferrovie dello Stato, the Italian State railways company.

[4] At the beginning of month t, the results for $t-2$ are released.

[5] For ease of exposition we avoid creating further acronyms and, from now on, we maintain the same names for the transformed variables. This should create no confusion.

Figure 1: The time series.

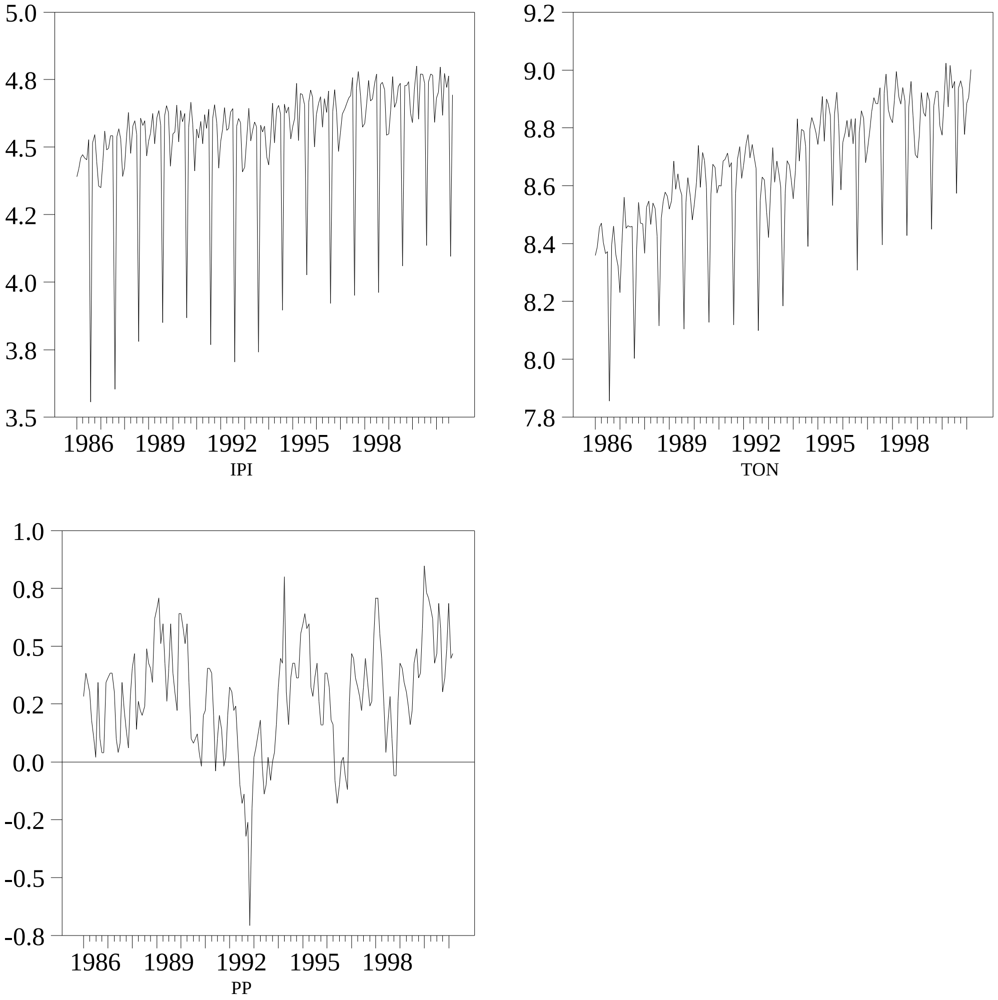


Figure 2: The spectra.

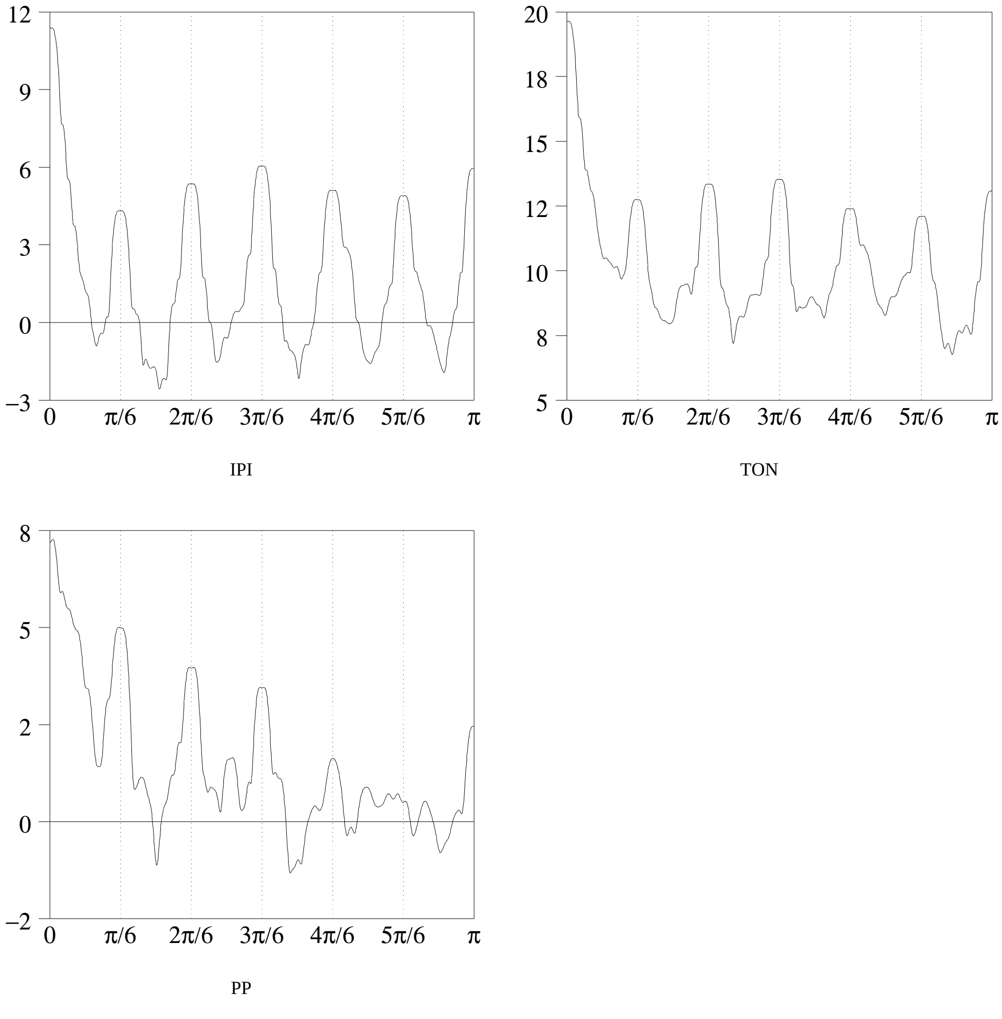


Table 1: Tests for unit roots

frequency	*IPI* (3 lags)		*TON* (3 lags)		*PP* (no lags)	
0	-2.64		-3.51	*	-2.68	
$\pi/6$	5.71		5.49		15.75	**
$\pi/3$	2.59		6.48		13.74	**
$\pi/2$	6.90	*	11.69	**	14.16	**
$2\pi/3$	12.00	**	8.74	*	14.16	**
$5\pi/6$	6.28		7.55	*	16.48	**
π	-1.90		-1.76		-2.45	

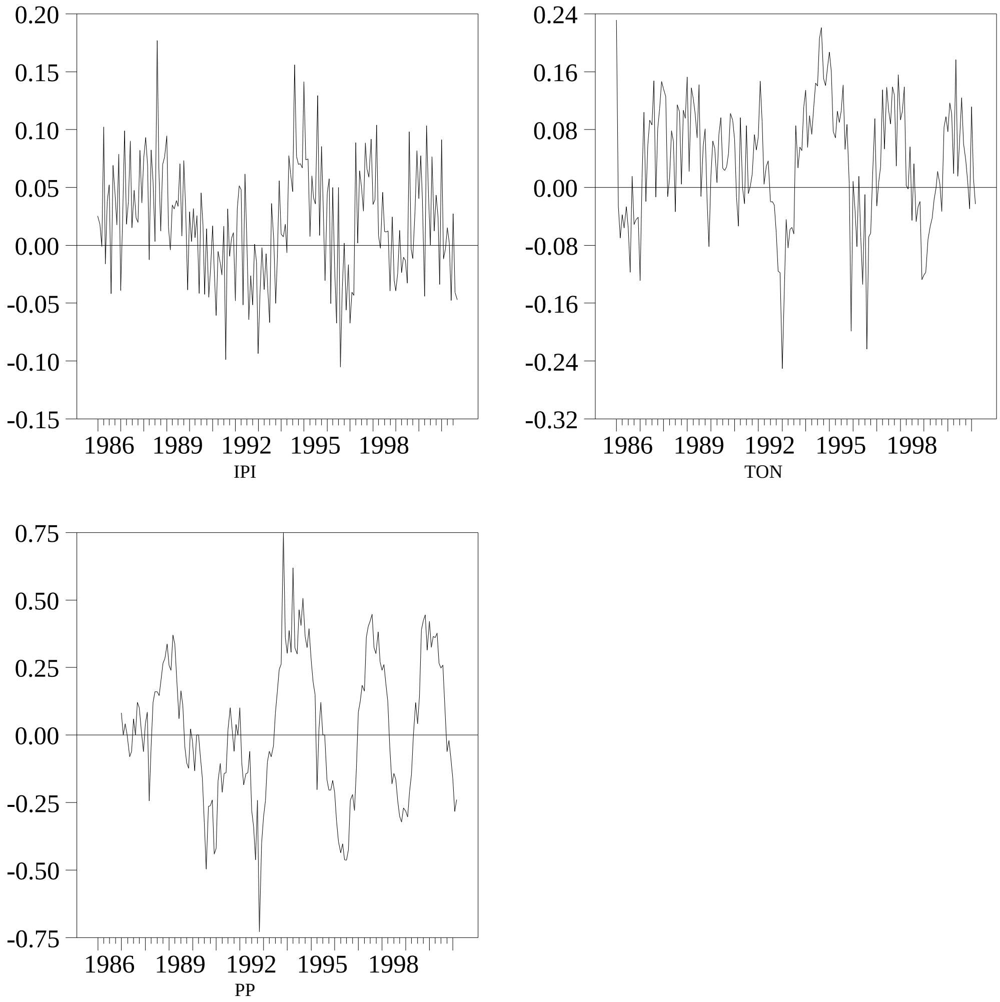
t-tests for the 0 and π frequencies, F-tests for the others. Values significant at 5% and 1% are indicated by '*' and '**', respectively.

The relative importance of trend, seasonality and cyclical movements can be better appreciated by resorting to the Fourier representation of time series. An estimate of the spectrum[6] of the series is showed in Figure 2. This confirms the previous observations, showing a similarity between the spectra of *IPI* and *TON*, with a concentration of power at low and seasonal frequencies. A slight different pattern emerges when *PP* is considered: its long-run component has a smaller peak in the spectrum, while considerable power seems to be present at the cyclical frequencies. Important peaks are present, moreover, at the fundamental seasonal frequency and, though less important, at frequencies associated with periodicities of 4 and 6 months.

2.2 Stochastic properties

The stochastic properties of the three series can be examined in a more formal manner, testing for the presence of unit roots. Given the nature of the data we are dealing with, it is natural test for such roots at the zero and at the seasonal frequencies. The framework retained in this paper to this aim is the one detailed in Beaulieu and Miron (1993). The test regression includes a constant, a trend, eleven seasonal dummies, and the number of lags of the dependent variable sufficient to whiten the residuals. The results reported in Table 1 reject the presence of a unit

Figure 3: Differenced series.

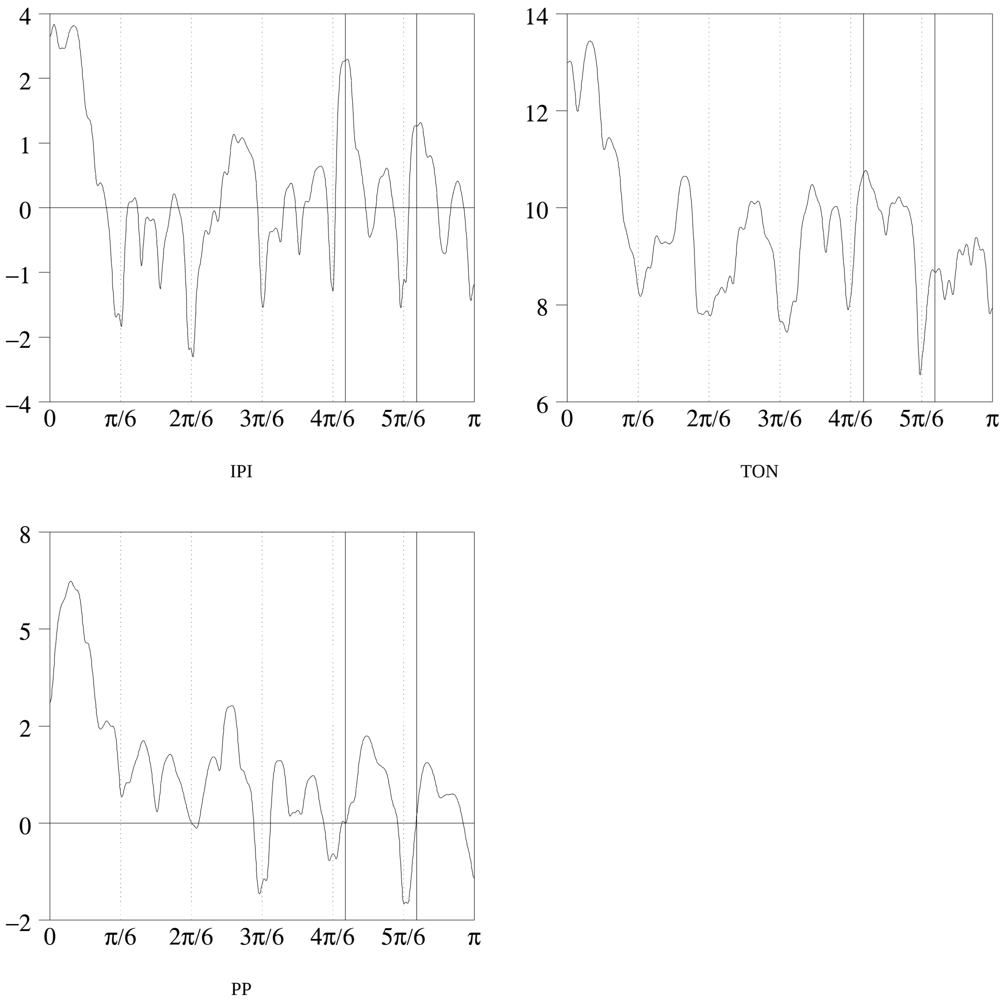


root at frequency zero for *TON*, while it is not rejected for the others. In all cases the presence of all the twelve unit roots is strongly rejected.

The conclusion we draw from the previous evidence is that the presence of unit roots at some seasonal frequencies cannot be overall excluded. The application of the seasonal difference operator produces the series plotted in Figure 3, where the cyclical pattern emerges more clearly, especially for *PP*, while *IPI* and *TON* are characterized also by strong irregular movements, some of which can be reasonably attributed to trading days effects. This feature is even clearer if we consider the spectral representation depicted in Figure 4, where seasonal peaks have disap-

[6] Actually, the spectral density is correctly defined only for stationary time series; in particular, when some unit roots are present in correspondence of certain frequencies, the usual expression for the spectral density would take the value $+\infty$ at those frequencies; nevertheless, discarding them the so called *pseudo spectrum* (Bell, 1984) can be considered. In our case the spectrum has been estimated by smoothing the periodogram using a rectangular spectral window. A cosine taper has been applied to the data. The spectral bandwidth is 0.048π.

Figure 4: Spectra of the differenced series.

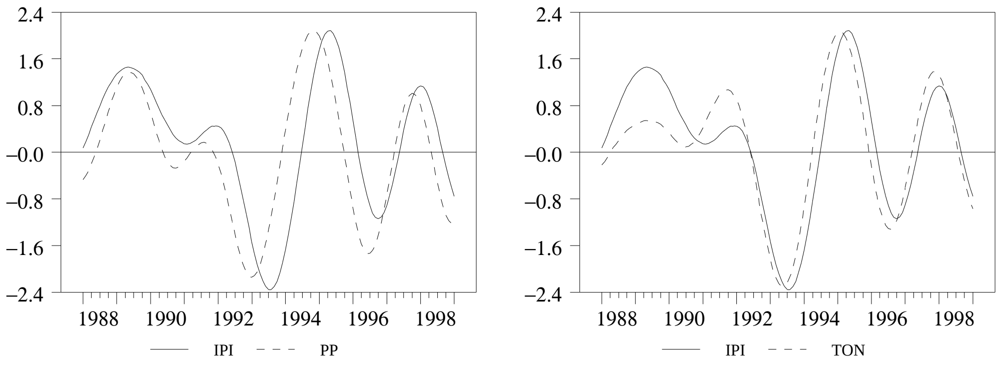


peared and the long term component at zero frequency is less pronounced. This makes much more evident the presence of a trading days pattern for the series *IPI* and *TON*, represented by peaks in the spectrum at the frequencies highlighted by vertical lines in the figure (on this aspect see Cleveland and Devlin, 1982).

2.3 Cyclical characteristics

One important characteristic of the Italian industrial production index is represented by its strongly cyclical behavior. Indeed, this is perhaps the most important feature one is normally interested in when formulating the forecasts. Moreover, early detection of cyclical up- and downswings is extremely useful as long as one is interested in analyzing the business cycle or when further smoothing is to be carried out (*e.g.* for seasonal adjustment or trend extraction) because, in this latter case, correct prediction of a turning point can reduce dramatically the extent of revisions implied by the smoothing itself. This is the reason why series used to

Figure 5: Cyclical components (2-8 year period component, standarized).



help forecasting industrial production index should be characterized by a regular lead on the latter at cyclical frequencies. We show that this is the case for the series considered in this paper.

We extract the cyclical component of each series by means of the *band-pass* filter developed by Baxter and King (1999). The estimated components are showed plotted in Figure 5 over the period 1988:1 -1999:1, leaving out two years of observation at the beginning and at the end of the available sample, since the band-pass filter is a symmetric one and the filtered observations at the extremes cannot be estimated. The figure shows a clear and regular lead of the production prospects over the industrial production, which is consistent with the nature of this series. Less clear from the figure is the leading nature of the *TON* series, which nevertheless holds, as highlighted in Table 2. The cyclical component of *PP* is confirmed to lead consistently that of *IPI*, on average by 5 months, with a high correlation. The correlation is even more pronounced if *TON* is considered, even though with a shorter lead (2 months on average).

Table 2: Correlation of the cyclical components of TON and PP with that of IPI

	$\rho(0)$	$\rho_{\max}$	lead (+) or lag (-)
TON	0.87	0.92	+2
PP	0.68	0.90	+5

”$\rho(0)$” is the correlation between the series and IPI; ”$\rho_{\max}$” is the maximum cross-correlation; ”lead” is the time interval (in months) at which the maximum cross-correlation is observed.

3. THE FORECASTING MODEL

An explicit goal of this study is that of finding a reliable but simple model to forecast the monthly Italian industrial production index. On the one hand, empirical evidence on the forecasting performance of nonlinear models is mixed (see *e.g.* Clements and Krolzig, 1998; Huh, 1998; Marchetti and Parigi, 2000; Simpson *et al.*, 2000). Franses and van Dijk (2001) suggest that linear models with simple seasonal components offer advantages over more complicated ones in terms of their short-term forecasting accuracy. On the other hand, we feel that the single-equation framework often used to forecast the industrial production index (see *e.g.* Marchetti and Parigi, 2000; Simpson *et al.*, 2000) offers an oversimplified option and does not allow for multi-step dynamic forecasts. For all these reasons our investigation rests on the well established VAR framework.

Given that we use seasonal time series, an aspect that deserves special attention is the parameterization of the VAR. According to the results listed in Section 2, the three time series that we consider have different seasonal properties: all display the presence of at least one unit root, but none of them seems to possess all the seasonal roots equal to unity. This implies that if we parameterize the VAR in seasonal differences, we are likely to over-difference the series. However, we believe that unit-roots pre-testing is useful for forecasting, despite the potentially low power of the tests (see Diebold and Kilian, 2000, for a discussion related to unit roots at the zero-frequency). Indeed, not much is known about the effects on forecasting performance deriving from imposing all the seasonal roots at unity when this is not the case in reality. To the best of our knowledge, the empirical evidence does not offer a definitive answer, though there are indications that filtering out only the correct unit roots in general does not produce superior forecasts (see *e.g.* Clements and Hendry, 1997; Gustavsson and Nordström, 1999; Lyhagen and Löf, 2001; Paap *et al.*, 1997). In particular, Lyhagen and Löf (2001) suggest that when the model is not known and the aim of the modeling 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. Therefore, given also that we deal with monthly data, we parameterize our VAR in seasonal differences. Also note that, given the standard short-term economic analysis practices, we are mostly interested in forecasting the annual growth rates.

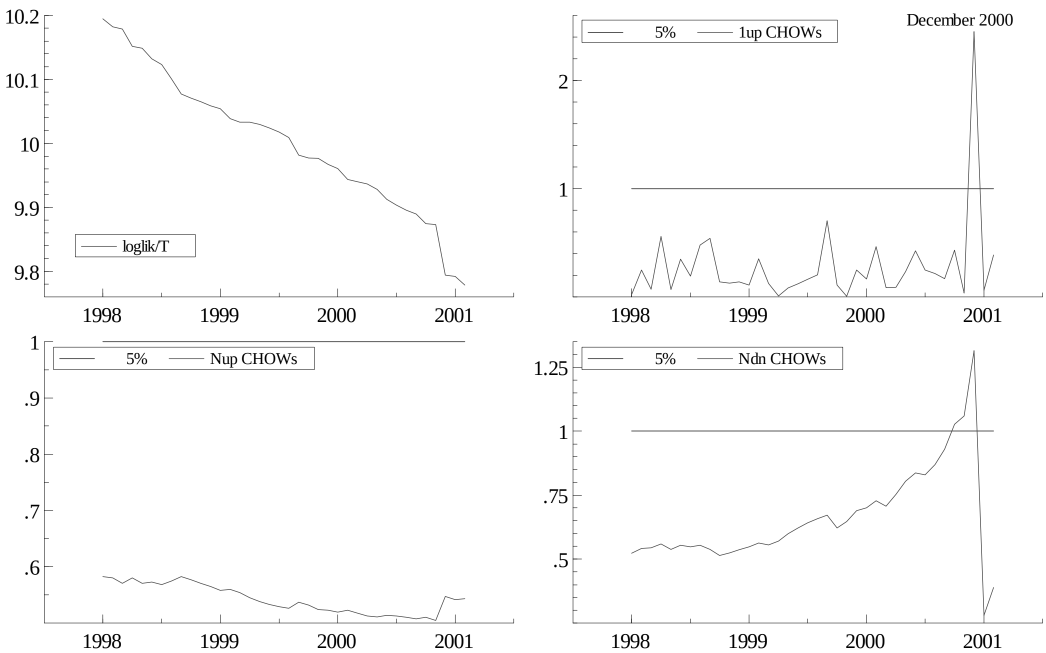
As far as model selection is concerned, we rely on the general-to-specific approach and we start from a fairly heavily parameterized VAR. Though seasonal differences do effectively filter out the seasonal components of the series (we do not need to use seasonal dummies), nevertheless they still show high and slowly decreasing

Table 3: Main VAR diagnostics: estimation period 1988.3-1997:12

	σ	Corr(Act., Fit.)	AR 1-12	Norm.
$\Delta\Delta_{12}IPI$	0.020	0.962	0.355	0.310
$\Delta\Delta_{12}TON$	0.045	0.868	0.283	0.620
$\Delta\Delta_{12}PP$	0.110	0.761	0.217	0.418
VAR			0.221	0.601
Parameter constancy forecast tests (1998:1-2001:2)				
F_{Ω}	0.524			
$\mathsf{F}_{\mathsf{V}(\mathbf{e})}$	0.944			
$\mathsf{F}_{\mathsf{V}(\mathbf{E})}$	0.959			

The Table reports the standard error of each equation in the VAR (σ), the correlation of actual and fitted values (Corr(Act., Fit.)), 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 residuals normality (Norm.). The p-values of the tests on the residuals of the VAR as a whole are also reported in the row labelled "VAR". The values reported for the parameter constancy forecast tests are p-values of the tests in their F-form. The first one (F_{Ω}) does not consider parameter uncertainty.

Figure 6: VAR Chow tests, 1998:1-2001:2



autocorrelations which make it difficult to find a valid (subset) reduction of the starting model (see also Krolzig, 2001). For this reason we reparameterize the VAR as a VECM in seasonal differences. Indeed, this proves useful to obtain quasi-orthogonal regressors. The starting unrestricted model takes the form

$$\Delta\Delta_{12}\mathbf{y}_t = \beta'\Delta_{12}\mathbf{y}_{t-1} + \sum_{j=1}^{13}\gamma_j'\Delta\Delta_{12}\mathbf{y}_{t-j} + \phi'\mathbf{d}_t + \boldsymbol{\varepsilon}_t \tag{2}$$

where $\Delta = (1 - L)$, $\Delta_{12} = (1 - L^{12})$, L is the usual lag operator such that $L^p z_t = z_{t-p}$, $\mathbf{y}_t = (IPI_t, TON_t, PP_t)'$, and $\mathbf{d}_t$ are the deterministic components. We find successful not to include all the seasonal dummies in the model (the p-value of the test for the exclusion of all the seasonal dummies except that for August is 0.991). Rather, $\mathbf{d}_t$ includes, besides the constant and the dummy for August, three specific impulse dummies, and two special dummies for August and December that take the value 1 when production prospects (PP) are positive and -1 when they are negative.[7] This approach is justified on the grounds that it is common practice for firms in Italy to adjust production to demand by prolonging (shortening) summer and Christmas holidays when demand is low (high). Furthermore $\mathbf{d}_t$ includes also $\Delta_{12}\log(TD_t)$ and $\Delta_{12}\log(TD_{t-1})$, with TD_t the number of trading days in month t. As is well known, the number of trading days significantly influences manufacturing activity. While the use of $\Delta_{12}\log(TD_t)$ is common in models for industrial production, the insertion of $\Delta_{12}\log(TD_{t-1})$ is fairly non-standard. However, in the presence of particularly unfavorable (favorable) trading days configurations, it is legitimate to expect that firms tend to compensate lower (higher) realized production in the following month. Indeed, the estimated coefficients of $\Delta_{12}\log(TD_t)$ and $\Delta_{12}\log(TD_{t-1})$ in our VAR are both highly significant and seem to confirm this view.

The VAR is sequentially simplified to obtain a more parsimonious parameterization. Even if the (subset) restricted VAR is more parsimonious than the starting one, nevertheless it is still rather highly parameterized including lags from 1 to 5, lag 9, and lags from 12 to 13. The p-value of the reduction is 0.8026, which indicates that no significant information is lost in the sequential simplification process. The main statistics and diagnostics of the VAR estimated over the period 1988:1-1997:12 are reported in Table 3.[8] The tests for parameter constancy, calculated over

[7] Strictly speaking, the use of these dummies is such that the model is no longer a VAR. However, given the rather special role of these variables, we prefer to continue denoting our model as "VAR". We used also a parameterization in which positive and negative dummies were separated, but the attached coefficients resulted not significantly different in absolute value. Therefore we preferred the more compact form described in the text.

[8] The results have been obtained using Pc-Fiml 9.30 (see Doornik and Hendry, 2000).

Table 4: Main VAR diagnostics: estimation period 1988.3-2001:2

	σ	Corr(Act., Fit.)	AR 1-12	Norm.
$\Delta\Delta_{12}IPI$	0.019	0.962	0.103	0.097
$\Delta\Delta_{12}TON$	0.043	0.865	0.056	0.774
$\Delta\Delta_{12}PP$	0.106	0.706	0.254	0.258
VAR			0.286	0.268

The Table reports the standard error of each equation in the VAR (σ), the correlation of actual and fitted values (Corr(Act., Fit.)), 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 residuals normality (Norm.). The p-values of the tests on the residuals of the VAR as a whole are also reported in the row labelled "VAR".

the forecast evaluation sample (see next section), do not reject structural stability. The same conclusion is reached looking at Figure 6, even if from the inspection of the graphics reported there the indication that an impulse dummy variable would have been advisable for December 2000 arises. Indeed, it should be stressed that, when first published, most analysts considered that figure as largely unexpected.

For completeness, in Table 4 we report the main statistics and diagnostics of the VAR estimated over the full sample (1988:3-2001:2).

The final model we use to actually produce forecasts is further simplified by eliminating non significant deterministic elements from individual equations.[9]

4. FORECAST EVALUATION

In this section we evaluate the forecasting ability of our VAR as opposed to an ARIMA model and to the forecasts released by CSC and IRS over a fairly long period (1998:1-2001:2).[10] Given that we are especially interested in forecasting industrial production annual growth rates, all forecasts comparisons refer to this variable. To make the evaluation more interesting, the ARIMA model is enriched

[9] This further simplification and the procedure to routinely produce the forecasts are implemented in WinRATS 5.00 (see Doan, 2000).

[10] Time series of past CSC forecasts have been obtained from the Confindustria Website (`http://www.confindustria.it/DBImages.nsf/HTMLPages/Centro+studi`), and start from 1998:3. IRS forecasts have been retrieved from the articles published in the financial newspaper *Il Sole 24 Ore*.

with a deterministic part that includes trading days and Easter effects. This model is estimated recursively by maximum likelihood and the forecasts are produced using TRAMO (see Gómez and Maravall, 1998). Such an ARIMA constitutes a very robust benchmark to beat.

Perfectly fair forecasts comparisons would require the use of homogeneous forecasting criteria among the competing models (see *e.g.* Tashman, 2000). However, we want to compare our forecasts with those from a model of which we don't know many details, and even with those derived from a survey. For this reason we believe that, while perfectly homogeneous conditions are essential when comparing the forecasting performance of alternative methods, they cannot be imposed when comparing real world forecasts. However, to increase comparability both the ARIMA and the VAR forecasts are based on a recursive scheme. Parameters are estimated with data ranging from 1 to t_0, and forecasts are produced for $t_0 + 1$, $\ldots$, $t_0 + n$ ($n \geq 1$); then parameters are estimated on the sample ranging from 1 to $t_0 + 1$, and forecasts are produced for $t_0 + 2$, $\ldots$, $t_0 + n + 1$, and so forth. In our application the forecast evaluation sample runs from January 1998 to February 2001: the estimation sample is adjusted in such a way that for each forecasting horizon we have 38 out-of-sample observations.

Comparisons are somewhat complicated by some peculiarities in the CSC and IRS forecast samples. In fact, CSC does not produce one-step ahead estimates for the month of July and two-step ahead projections for the month of August of each year. IRS does not release two-step ahead forecasts for the month of August of each year; additionally, we could not retrieve IRS forecasts for a couple of dates.[11]

Macroeconomic analysts might be interested, more than on the numerical indications arising from the forecasts, on their signs, since these can be perceived as warnings of expansions or contractions. For this reason we think that it is useful to start the investigation of the forecasting performance of our VAR model from an analysis of the directional forecasts. In our forecast sample there are 23 observations for which the industrial production annual growth rates are positive, 14 for which they are negative and one (July 2000) in which the annual growth rate is zero. We assume that if a prediction has wrong sign, but the difference with the actual growth rate is less than one percentage point, the sign of the forecast is correct. This avoids considering as wrong results close to zero.[12] The results from this comparison are reported in Table 5 and indicate that the gain with respect to

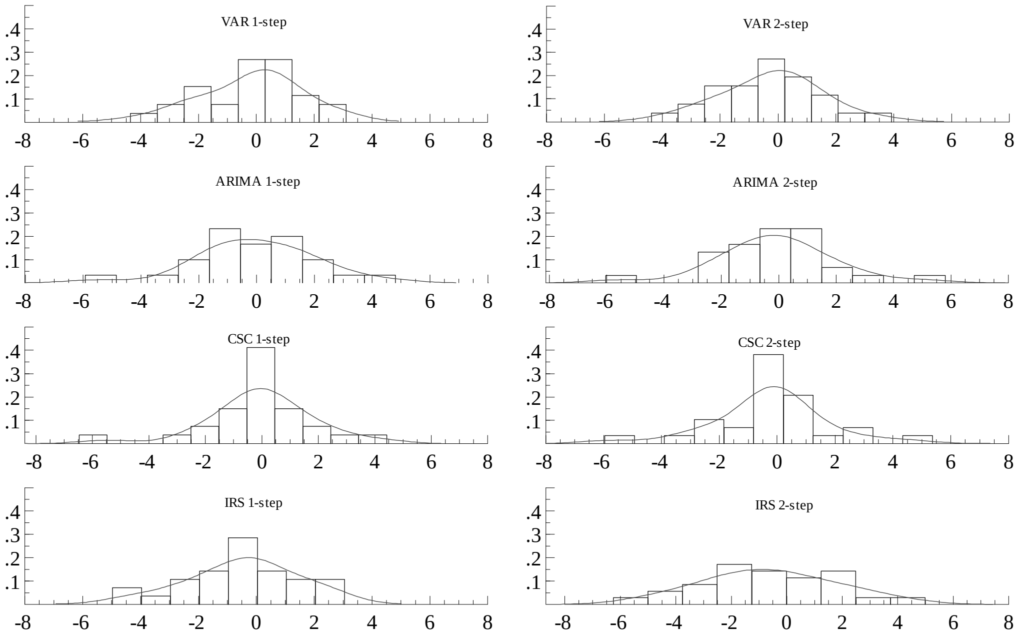
[11] This happened for example in corrispondence of dates for which IRS released only the seasonally adjusted figures.
[12] However, results do not change qualitatively if we use a strict criterion.

Table 5: Directional forecast errors: wrong predictions as percentage of valid observations.

steps ahead	1	2	3	6	12
ARIMA	13.16	13.16	7.89	31.58	13.16
VAR	13.16	13.16	13.16	10.53	10.53
NAIVE	45.95	41.67			
CSC	15.15	12.12			
$\text{ARIMA}_{\text{CSC}}$	12.12	15.15			
VAR_{CSC}	12.12	12.12			
IRS	21.62	23.53			
$\text{ARIMA}_{\text{IRS}}$	13.51	14.71			
VAR_{IRS}	13.51	11.76			

Wrong directional forecasts as percentage of valid observations. The one-step and two-step ahead "NAIVE" directional forecasts are given by sign($\Delta_{12}IPI_{t-1}$) and sign($\Delta_{12}IPI_{t-2}$), respectively. "ARIMA" and "VAR" with the subsctript "CSC" and "IRS" denote the statistics calculated on the forecasts from the ARIMA benchmark and the VAR model over same sample used for the CSC and IRS forecasts, respectively.

Figure 7: Estimated densities of forecasting errors.



a naive forecasts defined as $\mathrm{sign}(\widehat{\Delta_{12}IPI_t}) = \mathrm{sign}(\Delta_{12}IPI_{t-i})$ with $i = (1, 2)$ is substantial. Furthermore, the ARIMA, CSC, and VAR projections show similar directional errors, while IRS prediction errors nearly double the others. Finally, the ARIMA benchmark shows a large fraction of errors corresponding to the six-step ahead forecasts. It should be noted, however, that contingency tables-based tests on the directional forecasts[13] are always very significant, indicating that direction-of-change forecasts are informative for all the predictions considered in this paper. This remains true even for the forecasts of $\Delta\Delta_{12}IPI_t$.

Figure 7 plots the estimated densities of the forecasting errors for the different models, computed on the same common sample. Note that IRS forecasts seem to be the most uncertain, while CSC projections are the most concentrated around zero, even if they have fatter tails than the VAR predictions. Though informative, Figure 7 could be not very significant on statistical grounds, given that the sample is rather short. The features of the forecasts in terms of their mean absolute error (MAE) and mean error (ME) are reported in Table 6. Our VAR model's forecasts uniformly outperform the others in terms of MAE. Furthermore, note that the ratio of the twelve-step ahead to the one-step ahead MAE is 2.02 for the ARIMA model and only 1.19 for the VAR. This shows how important is the cyclical information embodied in *PP* and *TON*.

[13] The detailed results are not reported for brevity. On the characteristics of the tests see *e.g.* Diebold and Lopez (1996).

Table 6: Mean absolute forecasting errors and mean forecasting errors of yearly growth rates (percent) forecasts of industrial production

steps ahead	1	2	3	6	12
ARIMA	1.64	1.62	1.65	2.16	3.31
	0.16	0.17	0.16	0.23	-0.26
VAR	1.25	1.27	1.31	1.54	1.49
	0.11	0.17	0.22	0.18	0.18
CSC	1.33	1.26			
	-0.12	0.32			
$\text{ARIMA}_{\text{CSC}}$	1.59	1.31			
	0.05	0.01			
VAR_{CSC}	1.23	1.21			
	0.03	0.13			
IRS	1.46	1.94			
	0.32	0.50			
$\text{ARIMA}_{\text{IRS}}$	1.56	1.32			
	0.12	0.07			
VAR_{IRS}	1.29	1.24			
	0.11	0.20			

For each number of steps ahead, the first row reports the mean absolute error (MAE), while the second one shows the mean error (ME). ”ARIMA” and ”VAR” with the subsctript ”CSC” and ”IRS” denote the statistics calculated on the forecasts from the ARIMA benchmark and the VAR model over same sample used for the CSC and IRS forecasts, respectively.

In order to get a better assessment of the relative forecasting ability of our VAR as opposed to the other forecasts, we perform formal tests of (pairwise) equal forecasting performance and forecast encompassing. Both classes of tests are variants of the test for predictive accuracy proposed by Diebold and Mariano (1995).[14] 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 (or survey) i. Denote by $d_t = g(e_{it}) - g(e_{jt})$ with $g(\cdot)$ some arbitrary (non necessarily symmetrical) pre-specified function. The null hypothesis of equality of expected forecast performance is $\mathsf{E}(d_t) = 0$. It is natural to consider $\bar{d} = n^{-1}\sum_{t=1}^{n} d_t$, so that $\sqrt{n}(\bar{d} - \mu_d) \xrightarrow{\mathrm{d}} \mathsf{N}\left(0, 2\pi f_d(0)\right)$, where μ_d is the population mean of d_t and $f_d(0)$ is the spectral density of d_t at frequency zero, which in turn is $f_d(0) = (2\pi)^{-1}\sum_{\tau=-\infty}^{\infty}\gamma_d(\tau)$ with $\gamma_d(\tau)$ the lag-τ autocovariances. Diebold and Mariano (1995) propose basing the test of equal forecasting accuracy on

$$DM = \frac{\bar{d}}{\sqrt{n^{-1}2\pi\widehat{f_d(0)}}} \tag{3}$$

which, under the null, tends to $\mathsf{N}(0,1)$ 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.* (1997, 1998) propose modifying the test as

$$DM^* = \left(\frac{n + 1 - 2h + n^{-1}h(h-1)}{n}\right)^{1/2} DM \tag{4}$$

and comparing the results with the critical values from the Student's t distribution with $(n-1)$ degrees of freedom.

When comparing forecasting accuracy, in this paper we use $d_t = |e_{it}| - |e_{jt}|$: when performing tests of forecast encompassing, d_t becomes $d_t = e_{it}(e_{it} - e_{jt})$ (see Harvey *et al.*, 1998). Under the null, forecast i encompasses forecast j and $\mathsf{E}(d_t) = 0$: under the alternative, forecast i could be improved by incorporating some of the features present in forecast j.

In this paper we use the DM^* version of the tests. In order to obtain a consistent estimate of $f_d(0)$, we follow the recommendations contained in Diebold and Mariano (1995) and Harvey *et al.* (1997) and use an unweighted sum of the sample autocovariances up to $h-1$, that is $2\pi\widehat{f_d(0)} = \widehat{\gamma}_0 + 2\sum_{\tau=1}^{h-1}\widehat{\gamma}_\tau$, with $\widehat{\gamma}_k$ the lag-k sample autocovariance.

[14] The variants are those introduced by Harvey *et al.* (1997, 1998).

Table 7: Predictive accuracy tests

	1-step ahead forecasts			2-step ahead forecasts		
$M_i \setminus M_j$	VAR	CSC	IRS	VAR	CSC	IRS
ARIMA	1.306 (0.200)	1.325 (0.195)	0.800 (0.429)	1.193 (0.241)	0.981 (0.334)	-2.268 (0.030)
VAR		-0.342 (0.735)	-0.628 (0.534)		-0.165 (0.870)	-2.245 (0.032)
CSC			-0.383 (0.704)			-3.127 (0.004)
	ARIMA vs VAR (3 to 12 steps ahead)					
steps		3	4	5	6	7
		1.053 (0.299)	2.003 (0.053)	1.823 (0.076)	1.230 (0.226)	2.299 (0.027)
steps		8	9	10	11	12
		2.450 (0.019)	2.701 (0.010)	2.267 (0.029)	4.911 (0.000)	6.829 (0.000)

Modified Diebold-Mariano tests based on $d_t = |\hat{e}_{it}| - |\hat{e}_{jt}|$. The DM^* statistics and their p-values under the null (in brackets) are reported.

Two remarks are important at this stage. First, given that CSC and IRS forecasts present some missing values, in the computation of $2\pi\widehat{f_d(0)}$ we use (see Harvey, 1989, p.329; Robinson, 1985)

$$\widehat{\gamma}_k = \frac{\sum_{t=1}^{n-k}(d_t^\dagger - \bar{d})(d_{t+k}^\dagger - \bar{d})}{\sum_{t=1}^{n-k} a_t a_{t+k}} \tag{5}$$

where $d_t^\dagger$ is d_t with zeros replacing the missing values, and $a_t = 1$ when d_t is observed and $a_t = 0$ otherwise. Second, West (2001) demonstrates that when forecasts are based on estimated models and parameters estimation uncertainty is neglected, the forecast encompassing test tends to reject too often. This size distortion depends, among other things, on the number of out-of-sample forecasts used to compute the test. When the fraction n/t_0 is small, the distortion is likely to be small. In our case, $n/t_0 \approx 0.25$: this implies that a nominal 5% t test should slightly over-reject, but the actual size should not exceed 8%.[15] Given that correction of DM^* to take into account parameters uncertainty entails knowledge of both the models to be compared, we cannot in practice use the modifications suggested by West (2001).

In order to evaluate the forecasting accuracy of the various models, in Table 7 we report the results of the comparisons carried out on the different projections. The table shows that our model on average produces more precise forecasts than the others. However, the comparisons suggest that the difference is statistically significant only with respect to the two-step ahead forecasts released by IRS, and with the 4-to-12-step ahead ARIMA forecasts.[16]

The tests of forecast encompassing reported in Table 8 show less clear-cut results. From our viewpoint it seems relevant to note that, though the average quality of our forecasts is superior to that of IRS projections, nevertheless these embody some pieces of information that could improve both the one-step and the two-step ahead VAR forecasts. Note also that the converse apply even more strongly: in fact our VAR forecasts incorporate information that would be useful for improving IRS projections (and this, in the light of the previous findings, is an expected result). CSC predictions do not encompass ours. On the other hand, it is not entirely clear if our VAR forecasts encompass those elaborated by CSC: taking into account possible size distortions, the results seem to suggest that encompassing probably takes place for the one-step ahead predictions. The comparisons with the ARIMA indicate that the benchmark do possibly encompass the VAR forecasts only at the shortest horizon (one-step): on the contrary the predictions from the VAR encompass those from the ARIMA from the five-step ahead onward.

5. USING THE FORECASTS TO IMPROVE TREND-CYCLE REVISION PATTERNS

5.1 Characterization of the problem

This section illustrates one of the possible uses of the results of the forecasts derived from the model described in Section 3. While the main purpose of the model lies in the pure forecast of the raw industrial production index, nevertheless its results can be used, as already anticipated, to improve the construction of a cyclical indicator, reducing the revisions implied in its calculation.

[15] A nominal 3% should not exceed actual 5%. These computations follow West (2001, p.30) and are based on some rather unrealistic technical conditions. However, the values obtained in this way seem to act as upper bounds in the simulations carried out by West (2001, p.31).

[16] The test is significant (at the 10% significance level) for the 4- and 5-step ahead forecasts: it is not significant for the 6-step ahead forecasts (due to a single badly mispredicted value at 1999:12).

Table 8: Tests for forecast encompassing

		1-step ahead forecasts			
$\mathsf{M}_i \setminus \mathsf{M}_j$		ARIMA	VAR	CSC	IRS
ARIMA			1.805 (0.079)	1.395 (0.172)	1.979 0.056
VAR		2.246 (0.031)		1.832 (0.076)	2.745 (0.009)
CSC		1.188 (0.244)	2.945 (0.006)		1.475 (0.150)
IRS		2.085 (0.044)	3.603 (0.000)	1.524 (0.138)	
		2-step ahead forecasts			
ARIMA			2.088 (0.044)	3.460 (0.002)	3.222 (0.003)
VAR		1.999 (0.053)		2.094 (0.044)	2.405 (0.022)
CSC		2.084 (0.045)	2.822 (0.008)		-0.521 0.606
IRS		3.118 (0.004)	3.777 (0.000)	2.383 (0.023)	
		ARIMA vs VAR (3-12 steps)			
	3	4	5	6	7
	2.431 (0.020)	2.637 (0.012)	2.229 (0.032)	1.947 (0.059)	2.761 (0.009)
	8	9	10	11	12
	2.661 (0.014)	2.520 (0.016)	2.340 (0.025)	3.576 (0.000)	3.282 (0.002)
		VAR vs ARIMA (3-12 steps)			
	3	4	5	6	7
	2.148 (0.038)	2.089 (0.044)	0.960 (0.343)	1.189 (0.242)	0.737 (0.466)
	8	9	10	11	12
	0.622 (0.538)	1.164 (0.252)	1.660 (0.105)	0.762 (0.451)	0.197 (0.845)

Modified Diebold-Mariano tests based on $d_t = \hat{e}_{it}(\hat{e}_{it} - \hat{e}_{jt})$. The DM^* statistics and their p-values under the null (in brackets) are reported. The null hypothesis is that the forecasts produced by model M_i (column-wise) encompass those produced by model M_j (row-wise).

Let us consider the series IPI as composed by three (unobserved) components:

$$IPI_t = T_t + S_t + I_t. \tag{6}$$

The three elements are the trend (T_t), the seasonal (S_t) and the irregular component (I_t). The first represents the long term evolution of the series, together with oscillations associated with the business cycle. It should then be more correctly defined *trend-cycle*. Seasonal component represents movements which repeat themselves on a regular basis every year, while the irregular is a stationary, highly volatile and unpredictable component.

When looking for a cyclical indicator, one is normally interested in eliminating the seasonal and the irregular component, leaving only the trend-cycle. To estimate the latter, many criteria have been proposed. Some of them, such as the X-12-ARIMA seasonal adjustment procedure (Findley *et al.*, 1998), do not provide a statistical model for the components: others do provide an explicit characterization of the components. Among the latter, there are the structural time series approach (Harvey, 1989), and the ARIMA model-based approach (see Maravall, 1995, and the references therein).

The ARIMA model-based approach will be retained here, because of some appealing features. It is in fact quite simple to apply,[17] and produces a trend which is consistent with the seasonally adjusted figures officially provided by the Italian National Statistical Institute (ISTAT, 1999).

As pointed out before, the trend-cycle contains also movements associated to business cycle. In this paper we are not interested to further decompose long term trend and business cycle oscillations, also because trying to accomplish this task often produces results that can be contradictory and heavily dependent on the methodology used (Canova, 1998). In addition, the data we are dealing with feature frequent cycles, in the classical sense, so that no detrending is necessary in order for them to be more evident. Moreover, we are also interested in the absolute level of the series. In the end, if the forecasts produced by our model are useful to improve the construction of a trend-cycle indicator, they are likely to be equally useful with respect to the calculation of a purely cyclical indicator. The above considerations imply that whenever we refer to business cycle in this section, we do so in the sense of *classical cycle* and not in the sense of *growth cycle* (which consider detrended series).

[17] We apply it using the software TRAMO-SEATS (Gómez and Maravall, 1998).

5.2 Main features of the trend extracted by TRAMO-SEATS

The ARIMA model-based approach implies the identification and estimation of an ARIMA model for the observed series, with some possible deterministic components, such as trading days effects and outliers. ARIMA models for the components are then derived, using some identifying assumptions; among them there is the independence of the components in (6). Below, the main steps of the procedure are summarized.

Consider, for the aggregate monthly series y_t the following seasonal ARIMA representation:[18]

$$(1-L)\left(1-L^{12}\right)y_t = (1+\theta_1 L)\left(1+\theta_{12}L^{12}\right)\varepsilon_t \tag{7}$$

where L is the lag operator, and ε_t is i.i.d. $\sim \mathsf{N}(0,\sigma^2_\varepsilon)$. The autoregressive part of the model can be factorized as follows:

$$(1-L)\left(1-L^{12}\right) = (1-L)^2\,(1+L+L^2+...+L^{11}). \tag{8}$$

The first element in the right hand side of (8) implies two unit roots at the zero frequency of the spectral representation of y_t, while the second factor has eleven roots centered at the seasonal frequencies $2k\pi/12$, $k=1,2,...,11$. The first two unit roots, which are associated with the long term evolution of the series, can then be assigned to the trend component, while the other eleven are associated with the seasonal component. Restricting our attention to the trend, we obtain that it is given by the following ARIMA model:

$$(1-L)^2\,T_t = \theta_T\,(L)\,\varepsilon_{T,t} \tag{9}$$

with $\varepsilon_{T,t} \sim \mathsf{IN}(0,\sigma^2_{\varepsilon_T})$ and $\theta_T\,(L)$ a lag polynomial of order two. Nevertheless, the precise expression for the trend is not yet obtained, because there are many, possibly infinite, decompositions consistent with the aggregate model (7), which differ in the values of the coefficients in $\theta_T\,(L)$. The criteria used by TRAMO-SEATS to identify the components is to make them *canonical, i.e.* specifying them as free of noise as possible. The trend obtained is, therefore, the smoother trend among those obtainable from model (9) given the aggregate model (7). The canonical condition, together with the independence assumption, guarantees a unique decomposition.

[18] This example has been developed using the so called Airline model, which for monthly data is an ARIMA model $(011)(011)_{12}$. It has been chosen, both because it is the default model used by TRAMO-SEATS and it is the one actually used by ISTAT to seasonally adjust the IPI series.

Figure 8: Estimated trend component of the industrial production index.

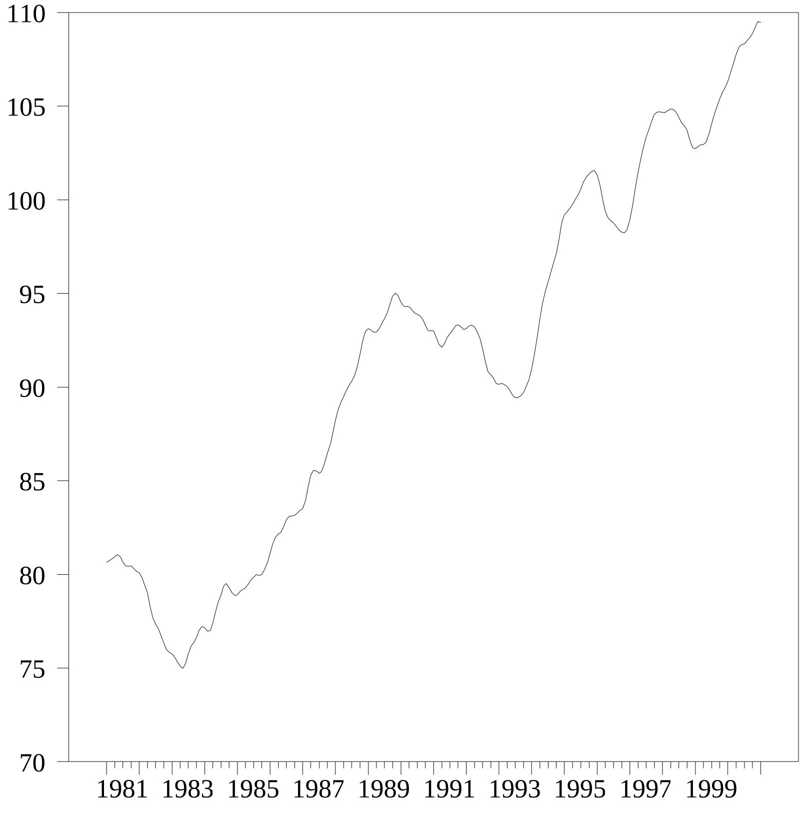


Figure 8 plots trend of the industrial production index estimated by TRAMO-SEATS. It clearly contains the long term evolution of the series, but also short term cyclical movements, up- and downturns.

Optimal (in a mean squared sense) estimation of the trend can be obtained using the Wiener-Kolmogorov filter (Maravall, 1995). Denote with $S(L) = (1 + L + L^2 + ... + L^{11})$, and with $\theta(L)$ the lag polynomial in (7). Considering an infinite realization of y_t, $Y = \{y_t\}_{t=-\infty}^{\infty}$, the trend estimator $\widehat{T_t}$ is given by the following expression:

$$\widehat{T_t} = \mathsf{E}(T_t|Y) = \frac{\sigma^2_{\varepsilon_T}}{\sigma^2_{\varepsilon}} \frac{\theta_T(L)\,\theta_T(L^{-1})S(L)\,S\left(L^{-1}\right)}{\theta(L)\,\theta(L^{-1})} y_t = \nu(L)y_t. \tag{10}$$

From equation (10) it is evident that the estimator of the trend consists in applying to the original series a symmetric, bidirectional, infinite filter $\nu(L)$. Moreover, invertibility of $\theta(L)$ ensures that the filter is convergent. This makes possible to render the procedure operational, approximating the infinite filter by truncation. Let now assume the truncated filter length is equal to $2r + 1$, so that r observations are lost at the end of the observed series. The usual solution is to extend the latter with the predictions coming out from the ARIMA model (7). This means that we actually have a sequence of preliminary estimates of the trend $\mathsf{E}(T_t|y_{t+i})$ ($i = 0, 1, ..r$) which gradually converge to the final one as long as predictions are replaced by the true values.

In this section we claim that the use of the forecasts coming out from the model described in Section 3 dramatically improves the preliminary estimate of the trend of IPI, thus making it a much better device in order to monitor the evolution of this variable. To justify the need for such an exercise, we rely on Bruno (2001), that shows that the revisions of the trend component of IPI can be unacceptably large. In particular, Bruno (2001) shows that the trend extracted by TRAMO-SEATS, while representing a good historical representation of the cyclical development of the industrial production index, is characterized by a deep worsening of its performance near the end of the series, which is the point economic and business analysts are mainly concerned with.

5.3 Revisions in the trend series

In order to check the importance of revisions in trend estimates, and to evaluate the advantages deriving from using our model's forecasts instead of the standard routine, we perform an historical simulation from January 1996 to February 2001, estimating the trend component with TRAMO-SEATS, for every period, for the original series and for series extended with 3, 6 and 12-step ahead forecasts.

The measure used to illustrate the revisions process is the following. Let $\widehat{T}_{t|t+k}$ be the estimate of the trend component at time t when a series of length $t+k$ ($k \geq 0$) is observed: the so called *concurrent estimate* is obtained when $k = 0$. The quantity

$$\hat{r}_k = \widehat{T}_{t|t+k} - \widehat{T}_{t|t+k-1} \qquad k = 1, 2, ... \tag{11}$$

represents, for every k, the monthly revision in the preliminary data, k months after the concurrent estimate.

We compute (11) for every month from January 1996 onward, obtaining a distribution of revisions for every k ranging from 1 (with 61 observations) to 61 (just one observation). In practice we are usually interested in, say, $k \leq 12$. We can therefore derive summary statistics of the monthly revisions: in particular it is interesting to check their variances, to see if the use of our model-based forecasts improves over the revision process. Figure 9 shows clearly how effective is the improvement in the revision pattern using the forecasts from our model. The black bar (labelled 'Original') is the variance associated with the monthly revision after k periods (x-axis) using the standard procedure: the variance behavior is characterized by a sharp decrease after the first five months, when it becomes negligible. The use of three-step ahead forecasts from our model reduces the variance of revision of about 35% during the first three periods, and of about 30% during the fourth and fifth month. Six-step ahead forecasts improve on this result, reducing by 50% the variance of revisions in the first three periods. Using 12-step ahead forecasts does not seem to give further significant gains.

5.4 Detection of turning points

In order to assess how important is the improvement in the revision process showed in the previous sub-section, it is possible to analyze if it helps, for example, in the detection of turning points. In order to do this we perform again a historical simulation, from 1996 onwards, applying a routine to detect the turning points in the trend estimated over the original series extended with 12 forecasts of TRAMO-SEATS, and with, respectively, 3, 6 and 12 forecasts coming from our model.[19] As a reference, we considered definitive turning points those identified with the observed series ending in February 2001.

[19] The need for extending the series derives from the fact that the turning points detection procedure discards the last five observations.

Figure 9: Variances of revisions of the trend.

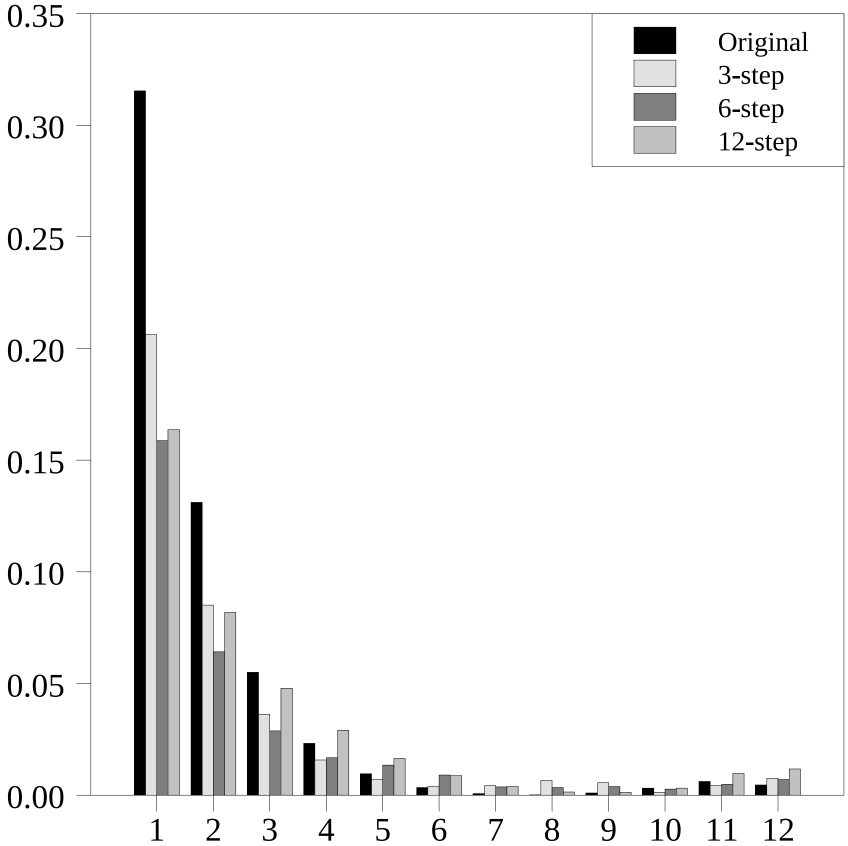


Table 9: Detection dates of turning points with different forecasts

Turning points	TRAMO-SEATS	3-step ahead	6-step ahead	12-step ahead
1995:12 (p)	96:09 (95:10)	96:05 (95:10)	96:03 (95:10)	96:02 (95:11)
1996:11 (t)	97:08 (96:12)	97:05 (96:12)	97:04 (96:12)	96:12 (96:12)
1998:04 (p)	99:02 (97:12)	98:10 (97:12)	98:07 (97:11)	98:04 (98:06)
1999:01 (t)	99:12 (99:01)	99:10 (99:01)	99:07 (99:01)	99:03 (99:04)
Mean lag	9.9	6.6	4.3	1.3

The table reports the dates of first detection of the turning points: ”p” denotes a peak, ”t” a trough. In the first column are listed the dates of the turning points as estimated using the whole time series up to February 2001. In parenthesis are reported the turning point locations as estimated at the detection date. Mean lag is the average lag of the detection.

It is important to stress that our aim is not to find out the best approach to signal a turning point, but simply to verify if the use of our forecasts helps in pursuing this objective. To assess this, we adopt the standard approach proposed by Bry and Boschan (1971).

The turning points identified by the procedure over the period 1996-2000 using the trend estimated over the actual data up to February 2001, are four, two peaks and two troughs: they are few, but going back further would have led to a too pronounced loss of data in order to estimate our model. The historical simulation is performed, again, reproducing as closely as possible a real world situation, that is re-estimating each month the model, leaving its structure unchanged. Table 9 shows the main results. The dates in the first column represent the turning points estimated as of February 2001, while the others are the months where the turning points were first detected. Dates in brackets represent the estimated locations of the turning points when first identified.

The mean lag in the detection of turning point with the original trend series is 10 months; this is true regardless the series being extended with 12 step-ahead TRAMO-SEATS predictions or not. The use of our model’s three-step forecasts improves the detection of the turning points in all cases, leading to a mean lag of 6.6 months. A further improvement is obtained with a longer forecasting horizon. With a six-step ahead forecast the turning point is detected, on average, after 4.3 months, while using a twelve step-ahead forecast it reduces to just 1.3 months.

Figure 10: Turning points detection under the standard procedure ('Concurrent') and using 12-step ahead VAR forecasts ('Forec.'). The date on the top of each panel indicates the last observation used in the trend estimation. The trend estimate based on actual data up to February 2001 ('Final') is also reported for comparison.

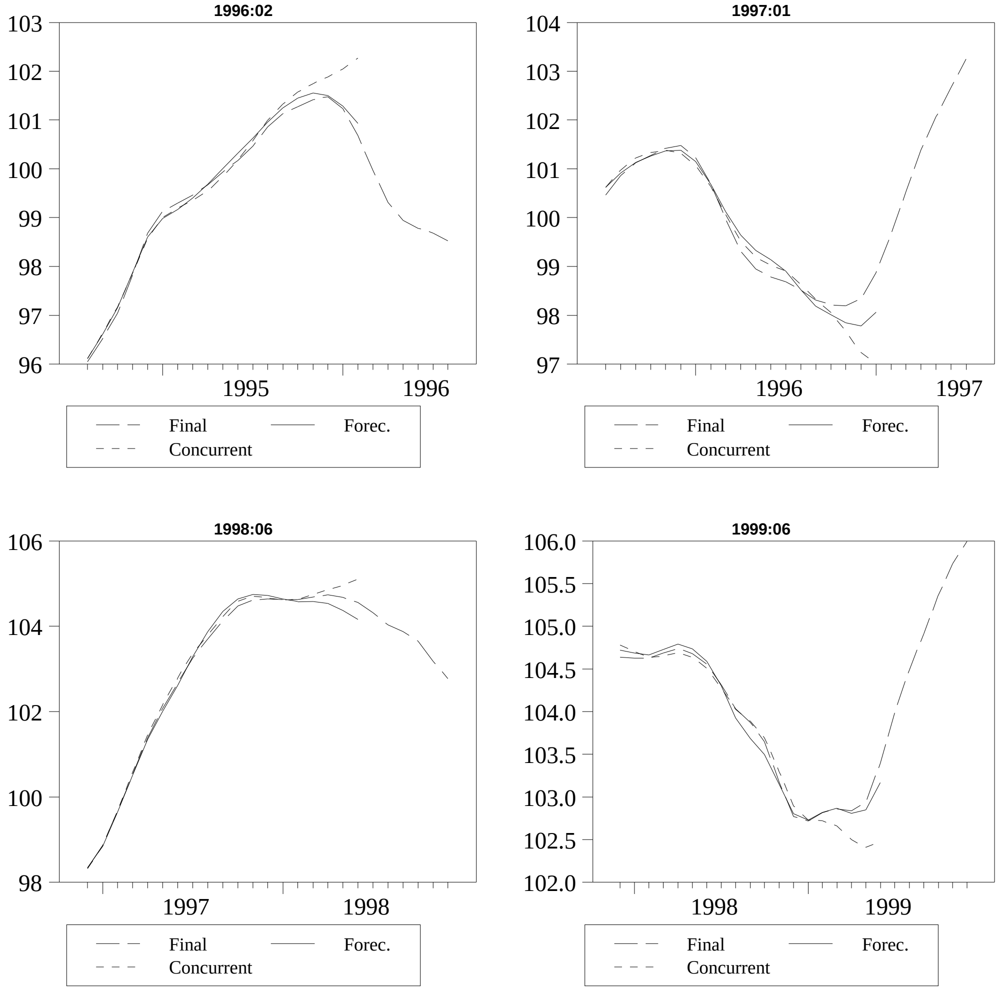


Figure 10 shows four cases in the neighborhood of the actual turning points where the performance of the trend obtained by extending the industrial production with 12-step forecasts of the VAR model (labelled 'Forec.' in the figure) is compared with the ordinary output of TRAMO-SEATS (labelled 'Concurrent'). The first appears to follow more closely the final estimate of the trend ('Final'), obtained using the observed time series as of February 2001. Visual inspection confirm very clearly the results illustrated in Table 9 and in Figure 9, that is the gain in precision in the trend estimate, in particular around turning points.

The procedure of Bry and Boschan in this context proves particularly robust against false signals, which never occur in our sample. Some problems emerge for the detection of the peak in 1998:4, which is sometimes located at the end of 1997; this is due to the "flatness" of the industrial production during that period. In addition, with the use of the twelve-step ahead forecast the turning point 1999:1 is identified the first time in January 1999, but not in the subsequent month: this is why in the table we report the value of March (from that month onward this turning point is always reported).

6. CONCLUDING REMARKS

In this paper we propose a simple VAR model to forecast Italian industrial production. We test for the predictive accuracy of our model over a fairly long forecast evaluation sample. We show that our VAR predictions outperform those produced on the basis of a robust ARIMA model, are on average at least as good as the survey-based projections elaborated by CSC, and more accurate than those deriving from the IRS econometric model. Furthermore, we show that using the VAR we are able to produce reliable forecasts on longer horizons. The forecast encompassing tests highlight that the different predictions embody different pieces of information that could be exploited to obtain even better forecasts. As long as one is interested only in forecasting horizons of at most two periods, this opens the possibility of investigating the opportunity of combining the forecasts: given that one of our goals is to produce multi-step dynamic forecasts, we do not pursue this route in the present paper.

We argue that obtaining good forecasts is essential to derive a reliable cyclical indicator using signal-extraction (smoothing) techniques. We show that this is the case by comparing the variance of revisions of a cyclical indicator estimated using our VAR's forecasts with that of the same indicator estimated using standard procedures: the information embodied in our predictions halves the uncertainty in the concurrent estimate of the cyclical indicator. This is also fundamental to timely detect turning points: the average gain in the delay with which a turning point is detected when using our forecasts is about nine months! We guess that a clear indication to praticitioners and economic analysts arise from these results: multi-step dynamic forecasts can improve substantially on the perception we can gain not only on the future, but also on the current phase of the economy.

REFERENCES

Altissimo, F., Marchetti, D.J. & Oneto, G.P. (1999). The Italian business cycle: New coincident and leading indicators and some stylised facts, ISAE, Documenti di lavoro, n.8/99 and 8A/99.

Baxter, M. & King, R.G. (1999). Measuring business cycles: Approximate band-pass filters for economic time series, *Review of Economics and Statistics*, 81, 575-593.

Bealieu, J.J. & Miron, J.A. (1993). Seasonal unit roots in aggregate U.S. data, *Journal of Econometrics*, 55, 305-328.

Bell, W.R. (1984). Signal extraction for nonstationary time series, *Annals of Statistics*, 12, 646-664.

Bodo, G., Golinelli, R. & Parigi, G. (2000). Forecasting industrial production in the Euro area, *Empirical Economics*, 25, 541-561.

Bovi, M., Lupi, C. & Pappalardo, C. (2000). Predicting GDP components using ISAE bridge equations econometric forecasting model (BEEF), ISAE, Documenti di lavoro n.13/00.

Bruno, G. (2001). Seasonal adjustment of Italian industrial production index using TRAMO-SEATS, ISAE, Documenti di lavoro n. 18/01.

Bry, G. & Boschan, C. (1971). Cyclical analysis of time series: Selected procedures and computer programs, NBER, Technical Paper n. 20.

Canova, F. (1998). Detrending and business cycle facts, *Journal of Monetary Economics,* 41, 475-512.

Clements, M.P. & Hendry, D.F. (1997). An empirical study of seasonal unit roots in forecasting, *International Journal of Forecasting*, 13, 341-355.

Clements, M.P., & Krolzig, H.-M. (1998). A comparison of the forecast performance of Markov-switching and threshold autoregressive models of US GNP, *Econometrics Journal*, 1, C47-C75.

Cleveland, W.S. & Devlin, S.J. (1982). Calendar effects in monthly time series: Modeling and adjustment, *Journal of the American Statistical Association*, 77, 520-528.

Diebold, F.X. & Kilian, L. (2000). Unit-root tests are useful for selecting forecasting models, *Journal of Business and Economic Statistics*, 18, 265-273.

Diebold, F.X. & Lopez, J.A. (1996). Forecast evaluation and combination, in G.S. Maddala and C.R. Rao (Eds.), *Handbook of Statistics*, Amsterdam: North-Holland.

Diebold, F.X. & Mariano, R.S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13, 253-263.

Doan, T.A. (2000). *RATS for Windows — Version 5*, Evanston: Estima.

Doornik, J.A. & Hendry, D.F. (2000). *Modelling Dynamic Systems Using PcGive, Vol II,* London: Timberlake Consultants Ltd.

Findley, D., Monsell, B., Bell, W., Otto, M. & Chen, B.C. (1998). New capabilities and methods of the X-12-ARIMA seasonal adjustment program, *Journal of Business and Economic Statistics,* 16, 127-152.

Franses, P.H. & van Dijk, D. (2001). The forecasting performance of various models for seasonality and nonlinearity for quarterly industrial production, Econometric Institute, Erasmus University Rotterdam, Econometric Institute Report EI 2001-14.

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

Gustavsson, P. & Nordström, J. (1999). The impact of seasonal unit roots and vector ARMA modeling on forecasting monthly tourism flows, FIEF Working Paper Series No. 150.

Harvey, A.C. (1989). *Forecasting, structural time series models and the Kalman filter*, Cambridge: Cambridge University Press.

Harvey, D.I., Leybourne, S.J. & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13, 281-291.

Harvey, D.I., Leybourne, S.J. & Newbold, P. (1998). Tests for forecast encompassing. *Journal of Business and Economic Statistics*, 16, 254-259.

Huh, C. (1998). Forecasting industrial production using models with business cycle asymmetry, *Federal Reserve Bank of San Francisco Economic Review*, n. 1, 29-41.

ISTAT (1999). La nuova strategia di destagionalizzazione degli indicatori congiunturali, *Note rapide*, n. 3.

Krolzig, H.-M. (2000). General-to-specific reductions of vector autoregressive processes, Department of Economics, University of Oxford, mimeo.

Lyhagen, J. & Löf, M. (2001). On seasonal error correction when the processes include different numbers of unit roots, SSE/EFI Working Paper Series in Economics and Finance No. 0418.

Maravall, A. (1995). Unobserved components in economic time series. In H. Pesaran et al. (Eds.), *The Handbook of Applied Econometrics*, vol. 1. Oxford: Basil Blackwell, pp. 12-72.

Marchetti, D.J. & Parigi, G. (2000). Energy consumption, survey data and the prediction of industrial production in Italy: A comparison and combination of cifferent models, *Journal of Forecasting*, 19, 419-440.

Paap, R., Franses, P.H. & Hoek, H. (1997). Mean shifts, unit roots and forecasting seasonal time series, *International Journal of Forecasting*, 13, 357-368.

Pappalardo, C. (1998). Una sintesi della letteratura sui modelli di previsione basati sull'utilizzo delle inchieste congiunturali, *Rassegna di lavori dell'ISCO*, n.3, 171-188.

Robinson, P.M. (1985). Testing for serial correlation in regression with missing observations. *Journal of the Royal Statistical Society*, Ser. B, 47, 429-437.

Tashman, L.J. (2000). Out-of-sample tests of forecasting accuracy: an analysis and review, *International Journal of Forecasting*, 16, 437-450.

Simpson, P.W., Osborn D.R. & Sensier, M. (2000). Forecasting UK industrial production over the business cycle, Centre for Growth and Business Cycle Research, School of Economic Studies, University of Manchester, mimeo.

West, K.D. (2001). Tests for forecast encompassing when forecasts depend on estimated regression parameters, *Journal of Business and Economic Statistics*, 19, 29-33.