Università degli Studi del Molise
Facoltà di Economia
Dipartimento di Scienze Economiche, Gestionali e Sociali
Via De Sanctis, I-86100 Campobasso (Italy)

ECONOMICS & STATISTICS DISCUSSION PAPER
No. 4/03

Forecasting Industrial Production and the Early Detection of Turning Points

by

Giancarlo Bruno
ISAE, Rome

and

Claudio Lupi
University of Molise, Department SEGeS

The Economics & Statistics Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. The views expressed in the papers are solely the responsibility of the authors.

Forecasting Industrial Production and the Early Detection of Turning Points[1]

Giancarlo Bruno[2]
Institute for Studies and Economic Analyses (ISAE)
Piazza dell'Indipendenza, 4. I-00185 Rome (Italy).
E-mail: g.bruno@isae.it
Tel. +39 06 4448 2716
Fax +39 06 4448 2249

Claudio Lupi
University of Molise, Faculty of Economics, Dept. S.E.G. e S.
Via De Sanctis, I-86100 Campobasso (Italy)
E-mail: lupi@unimol.it

April 11, 2003

[1]The first draft of this paper was mostly prepared while Claudio Lupi was working as a Research Director at ISAE. The present version of the paper is a revision of a preliminary draft that was circulated under the same title as ISAE working paper 20/01. The authors would like to thank two anonymous referees, audiences at ISAE, Dublin (22nd International Symposium on Forecasting), and Ente Einaudi (econometrics seminars) for comments. Fruitful discussions and suggestions from Gianluca Cubadda, Sergio de Nardis, John FitzGerald, Antonio García-Ferrer, Franco Peracchi, and Tommaso Proietti are gratefully acknowledged. None of them is responsible for any remaining error. 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.

[2]Corresponding Author.

Abstract

In this paper we propose a simple model to forecast industrial production in Italy up to 6 months ahead. We show that the forecasts produced using the model outperform some popular forecasts as well as those stemming from an ARIMA model used as a benchmark and those from some single equation alternative models. We show how the use of these forecasts can improve the estimate 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.

Keywords: Forecasting, VAR Models, Industrial production, Cyclical indicators.

JEL classification: C53, C32, E32.

1 Introduction

Forecasting the industrial production index is an important issue in short-term economic analysis. This is still true in contemporary developed economies where services are gaining 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. The interest in the topic is witnessed by the continuous effort devoted to investigating it (see, *e.g.*, Bodo *et al.*, 2000; Huh, 1998; Marchetti and Parigi, 2000; Osborn *et al.*, 1999). In addition, forecasts of industrial production can be useful in more general forecasting models. Finally, cyclical indicators of the manufacturing sector 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 useful in order to obtain reliable estimates of the most recent observations.

From a general standpoint it should be said that most of the existing models in Italy offer an early estimation (*nowcast*) of the industrial production, rather than a true forecast. Indeed, the official indicator is released by the Italian 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 highly reputed 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 an econometric 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 our model can well compete with the two aforementioned accredited forecasts in terms of predictive ability within the two-step ahead horizon, and offer reliable results up to six months ahead. In a spirit similar to McGuckin *et al.* (2001), we also show

[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 surveys. However, for brevity we will refer to the different forecasting devices as "models". This should cause no confusion.

how our projections can be used successfully to reduce uncertainty in the estimation of a cyclical indicator for the manufacturing sector derived from the industrial production series. This allows us also to improve substantially on the timely detection of turning points in the level of manufacturing activity. 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. The next Section presents the model: its forecasting ability is evaluated and compared with that of the other models in Section 3. The use of our forecasts to improve trend-cycle estimates and to reduce the delay with which turning points are detected, are discussed in detail in Section 4. The last Section concludes. Three appendices report some preliminary data analysis and estimation details.

2 The forecasting model

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 (*i.e.*, nowcasts). However, the goal of obtaining genuine 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 give priority, in the search of variables which will be used in the forecasting model, to those characterized by a leading pattern. 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] on future production prospects (PP) and the quantity of goods transported by railways (TON).[3] The first variable

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

[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`. 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

represents industrial entrepreneurs' opinions about future production. More precisely, the entrepreneurs are asked if they expect the production of their own firm to increase (+), to remain stable (=), or to decrease (–) in the following three-four months. The answers are conventionally synthesized using a *balance*, *i.e.* the share of "+" less the share of "–" answers. The variable (PP) obtained in this way is therefore bounded in the interval [–100, +100], and it is a natural candidate in a forecasting model of the industrial production index, given its timely availability, its explicit link with the variable to be forecast, and its expected lead with respect to 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.

A log transformation is used for the series IPI and TON while PP is rendered unbounded using the transform $-\log\left[200/(PP+100)-1\right]$.[4] None of the variables is seasonally adjusted.

Despite being extensively studied, nevertheless the Italian industrial production index is not an easy series to forecast, even considering seasonal differences (see Figure 1). An explicit goal of this study is that of finding a simple but reliable model to forecast the monthly Italian industrial production index up to six steps ahead.

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.*, 2001). 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.*, 2001) offers an oversimplified option that 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. Even if we test for unit roots in our series,[5] we adopt the rather practical strategy of modelling seasonal differences. The reason for this choice is threefold. First, following the standard short-term economic analysis practice we are mostly interested in forecasting

dello Stato, the Italian State railways company.

[4]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.

[5]The results are reported in appendix A.

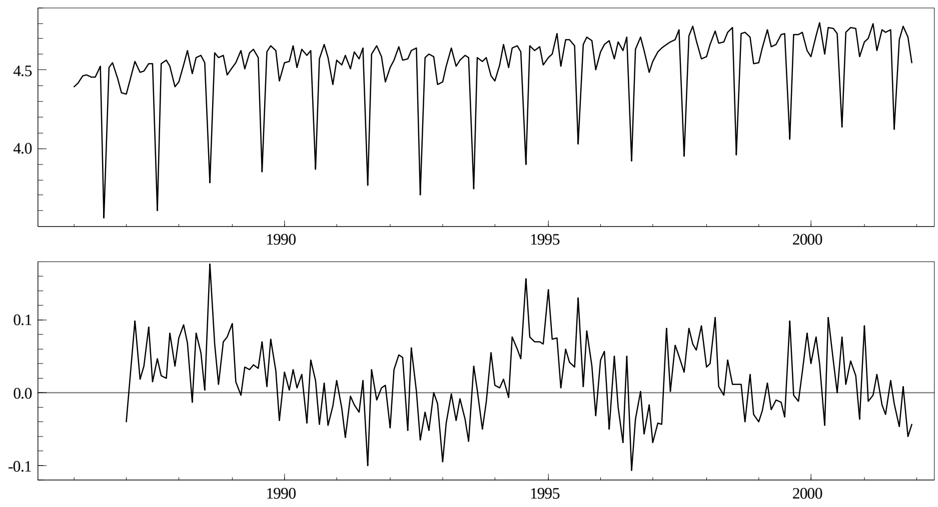

Figure 1: Log-industrial production index (above) and its seasonal differences (below).

the annual growth rates, not the levels of the series. Second, not much is known about the effects on forecasting performance deriving from imposing all the seasonal roots at unity when they are not all present in the observed time series: 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; Osborn *et al.*, 1999; 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. Last, but by no means least, differencing partially protects forecasting from sudden structural changes (Clements and Hendry, 1999), and we share the view that "[...] the current recommendation for the choice of a single model for forecasting seasonal economic time series up to a year ahead is that the annual difference specification should be the default choice" (Osborn, 2002, p.430).[6] The matter here is the balance between the reduction in forecast errors bias and the increase in variance: however this aspect can not be assessed *a priori*.

[6]Of course we are aware that, to the extent that annual differences induce non-invertibility, "protection" against breaks is somewhat reduced (see Osborn, 2002).

As far as model selection is concerned, we rely on the general-to-specific approach and we start from a fairly heavily parameterized VAR with 14 lags and some deterministic components. In particular, we simplify the model sequentially by excluding non significant lags, starting from the least significant, while checking at each simplification step the statistical properties of the residuals. Though seasonal differences do effectively filter out most of the seasonal component of the series, nevertheless they still show slowly decreasing autocorrelations which make it difficult to find a valid (subset) reduction of the starting model (see Krolzig, 2001). In order to obtain quasi-orthogonal regressors so to ease the reduction process, we reparameterize our unrestricted stationary VAR into the isomorphic form

$$\Delta\Delta_{12}\mathbf{y}_t = \boldsymbol{\beta}\Delta_{12}\mathbf{y}_{t-1} + \sum_{j=1}^{13}\boldsymbol{\gamma}_j\Delta\Delta_{12}\mathbf{y}_{t-j} + \boldsymbol{\phi}\mathbf{d}_t + \boldsymbol{\varepsilon}_t \tag{1}$$

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 that include, besides the constant and three specific impulse dummies (1991:08, 1995:08, and 1992:11) to correct for particularly large residuals,[7] a dichotomous variable that takes the value 1 in August when production prospects (PP) in July are positive and a similar variable when they are negative; other two dichotomous variables are specified along the same lines for December (having November as reference month for PP).[8] This approach represents an attempt to take into account possible interactions between seasonal variations and the business cycle in industrial production. These interactions can be theoretically justified by economic theory (see *e.g.* Cecchetti *et al.*, 1997), and can produce observable implications. Furthermore, it is very well known that in Italy it is common practice for the industrial firms to adjust production to demand by prolonging (shortening) summer and Christmas holidays when demand is low (high). We find successful not to include the seasonal dummies in the model (the p-value of the test for the exclusion of all the seasonals, except the special dummies for August and December, is 0.9999). Finally, $\mathbf{d}_t$ 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 widespread in models for

[7]The first two are related to the industrial production index, while the third is intended to correct an outlier in the production prospects (PP) series.

[8]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, for brevity we prefer to continue denoting our model as "VAR".

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: therefore we expect the parameters attached to $\Delta_{12}\log(TD_t)$ and $\Delta_{12}\log(TD_{t-1})$ to have opposite signs. Indeed, the estimated coefficients of $\Delta_{12}\log(TD_t)$ and $\Delta_{12}\log(TD_{t-1})$ in our VAR confirm our expectation and are both highly significant.[9]

The unrestricted 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 heavily parameterized including the lagged terms in annual differences and the lags from 1 to 5, lag 9, and lags from 12 to 13 for the terms in $\Delta\Delta_{12}$. The p-value of the reduction is 0.675, 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 1988:03-1998:12 are reported in Table 1.[10] The tests for parameter constancy, calculated over the forecast evaluation sample (1999:01-2001:12, see next section), do not reject structural stability.

The final model we use monthly to actually produce the forecasts is further simplified by eliminating non significant deterministic elements from individual equations.[11]

3 Forecast evaluation

In this section we evaluate the forecasting ability of our VAR as compared to an ARIMA model, to some simple but efficient single-equation alternatives, and to the forecasts released by CSC and IRS over a fairly long period (1999:01-2001:12).[12] Since our aim is to produce forecasts up to 6 months ahead, the forecasting performance comparison with CSC and IRS is reported for completeness and is justified on the grounds that we want to obtain a model that on the very short horizons behaves at least as well as those of these two honored forecasters. Given that we are especially interested in forecasting industrial production annual growth rates, all forecasts

[9]The coefficients (t-values) attached to $\Delta_{12}\log(TD_t)$ and $\Delta_{12}\log(TD_{t-1})$ in the equation for industrial production are 0.474 (8.75) and -0.271 (-3.63), respectively.

[10]The results have been obtained using Pc-Give 10.1 (see Doornik and Hendry, 2001).

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

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

Table 1: Main VAR diagnostics: estimation period 1988:03-1998:12

	σ	Corr(Act., Fit.)	AR 1-12	Norm.
$\Delta\Delta_{12}IPI$	0.020	0.959	0.051	0.217
$\Delta\Delta_{12}TON$	0.044	0.867	0.084	0.664
$\Delta\Delta_{12}PP$	0.106	0.754	0.090	0.478
VAR			0.093	0.485
Parameter constancy forecast tests (1999:01-2001:12)				
$\mathsf{F}_{\mathbf{\Omega}}$	0.143			
$\mathsf{F}_{\mathsf{V}(\mathbf{e})}$	0.569			
$\mathsf{F}_{\mathsf{V}(\mathbf{E})}$	0.607			

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 ($\mathsf{F}_{\mathbf{\Omega}}$) does not consider parameter uncertainty.

comparisons refer to this variable.

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 do not know many details (the IRS model), and even with those derived from a survey (CSC). For this reason we believe that, while perfectly homogeneous conditions are essential when comparing the forecasting performance of alternative *methods*, they cannot always be imposed when comparing *real world forecasts*. However, to increase comparability the ARIMA, the single-equation, and the VAR forecasts are all based on a common recursive scheme. Parameters are estimated with data ranging from 1 to t_0, and forecasts are produced for t_0+1, ..., t_0+n ($1 \leq n \leq 6$); then parameters are estimated on the sample ranging from 1 to t_0+1, and forecasts are produced for t_0+2, ..., t_0+n+1, and so forth.[13] In our application the forecast evaluation sample runs from January 1999 to December 2001: the estimation sample is adjusted in such a way that for each forecasting horizon we have $n = 36$ out-of-sample observations.

[13]For the single-equation models, the last observation available for estimation depends on the number of steps ahead to be forecast.

In order to evaluate the quality of our forecasts we consider different criteria. We first compute simple standard statistics (mean absolute forecasting error, root mean square error, mean error), and in order to draw some inference on the significance of the emerging differences we use the Diebold-Mariano (1995) approach as modified by Harvey *et al.* (1997, 1998). In particular, for the h-step ahead forecasts we use the modified Diebold-Mariano statistic:

$$DM^* = \left(\frac{n+1-2h+n^{-1}h(h-1)}{n}\right)^{1/2} \frac{\bar{d}}{\sqrt{n^{-1}2\pi \widehat{f_d(0)}}} \tag{2}$$

where $n = 36$, $\bar{d} = n^{-1}\sum_{t=1}^{n} d_t$, $d_t = g(e_{1t}) - g(e_{2t})$ with $g(e_{it})$ some arbitrary function of the forecasting errors from model $i\,\{\in 1,2\}$, and $\widehat{f_d(0)}$ is a consistent estimate of the zero-frequency spectral density of d_t.

When comparing forecasting accuracy, in this paper we use $d_t = |e_{1t}| - |e_{2t}|$: the null in this case is $\mathsf{E}(d_t) = 0$. This choice, as well as the emphasis on the MAE, is justified on the grounds that we would like to obtain a model that on average performs well, at the cost of some occasional relatively large error. When performing tests of forecast encompassing, d_t becomes $d_t = e_{1t}(e_{1t} - e_{2t})$ (see Harvey *et al.*, 1998): under the null, forecast 1 encompasses forecast 2 and $\mathsf{E}(d_t) = 0$; under the alternative, forecast 1 could be improved by incorporating some of the features present in forecast 2.

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.

Two remarks are important at this stage. First, given that CSC and IRS do not release their forecasts in summer according to different schemes, their projections present some missing values. 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.[14] To cope with this issue, in the computation of $2\pi\widehat{f_d(0)}$ we use (see Harvey, 1989, p.329; Robinson, 1985)

$$\widehat{\gamma_k} = \frac{\sum_{t=1}^{n-k}(d_t^\dagger - \bar{d})(d_{t+k}^\dagger - \bar{d})}{\sum_{t=1}^{n-k} a_t a_{t+k}} \tag{3}$$

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

[14] This happened in corrispondence of two dates for which IRS released only the seasonally adjusted forecasts.

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).

3.1 The ARIMA model

It is customary to compare the forecasts derived from a model to those relying only on univariate techniques. The reason for this, is that we want to be sure that using our model we can forecast at least as well as what we could do by exploiting only the information embodied in the variable of interest. Actually, the benchmark model we use in this paper is slightly more general than a "pure" ARIMA, being a regression model with ARMA errors, where the dependent variable is the stationary transform (seasonal difference) of IPI and the regressors are those included in $\mathbf{d}_t$ in equation (1) with the exception of the "special" dummies for August and December.

At any rate, for simplicity we will continue throughout the paper to denote this model as ARIMA. IPI has been considered both in the original scale and in logs. The BIC criterion is used to select the AR and MA polynomial orders, testing each lag combination up to 6 for both polynomials. An appropriate transformation of the likelihood (Findley *et al.*, 1998) allows to compare models both in levels and in logs. The selected model is an ARMA (2,2) on the seasonal difference of the untransformed IPI variable. The estimated AR polynomial has a pair of complex conjugate roots with a period of 37 months, which is consistent with the observed cyclical behavior of Italian industrial production.

This model has been estimated recursively by exact maximum likelihood[16] and 1 to 6 step-ahead forecasts have been produced. Such an ARIMA model

[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 software used to carry out the estimation is X-12-ARIMA ver. 0.2.9 (see U.S. Census Bureau, 2002).

Table 2: Forecasts evaluation statistics: VAR vs ARIMA.

Steps ahead		1	2	3	4	5	6
	ME	0.12	0.13	-0.04	-0.30	-0.11	-0.15
VAR	MAE	1.26	1.31	1.29	1.26	1.63	1.74
	RMSE	1.71	1.75	1.59	1.63	2.65	2.77
	ME	-0.01	-0.09	-0.08	-0.17	-0.26	-0.33
ARIMA	MAE	1.54	1.62	1.69	1.82	1.94	1.96
	RMSE	2.17	2.29	2.37	2.38	2.45	2.52
Predictive	VAR vs. ARIMA	-1.085	-1.122	-2.816	-3.504	-1.324	-0.831
accuracy	p-values	0.285	0.270	0.008	0.001	0.194	0.412
	VAR vs. ARIMA	1.228	1.250	0.778	0.303	1.342	2.149
Encompassing	p-values	0.228	0.220	0.442	0.764	0.188	0.039
	ARIMA vs. VAR	2.040	2.076	2.923	9.916	1.503	1.792
	p-values	0.049	0.045	0.006	0.000	0.142	0.082

Note. ME: Mean Error. MAE: Mean Absolute Error. RMSE: Root Mean Square Error.

constitutes a fairly robust benchmark to beat.[17]

The results, reported in Table 2, indicate that the MAE and RMSE of the VAR forecasts are uniformly smaller than those of the ARIMA, except for the five- and six-step ahead RMSE. The sudden increase of these two statistics in correspondence of five- and six-step ahead forecasts is due to one single large forecast error induced by the difficulty of predicting PP precisely at these longer horizons. The DM^* predictive accuracy tests denote a significant superiority of the VAR only for the three- and four-step ahead forecasts. However, the forecasts encompassing tests highlight that the VAR forecasts encompass the ARIMA ones at all horizons, with the exception of the six-step ahead forecasts. On the contrary, the ARIMA forecasts encompass those stemming from the VAR only for the five-step and, marginally, for the six-step ahead horizons.

3.2 Single-equation alternatives

While the univariate ARIMA model is a useful benchmark against which to test our proposed model, we also consider a set of single equation multivariate models which exploit as much information as possible from the same variables used by the VAR. In fact, at the half of month t, when we have to produce forecasts for months $t-1, t, t+1, ..., t+4$, the most recent observations on the variables that enter our model are $\{IPI_{t-2}, TONN_{t-1}, PP_{t-1}\}$. This fact is not exploited in the VAR, where we consider only the information set up to $t-2$ for all variables, in this way ignoring the available information for

[17] The same ARIMA is used also in the applications described in Section 4.

Table 3: Forecasts evaluation statistics: VAR vs Single-equation alternatives (SEQ).

Steps ahead		1	2	3	4	5	6
	ME	0.12	0.13	-0.04	-0.30	-0.11	-0.15
VAR	MAE	1.26	1.31	1.29	1.26	1.63	1.74
	RMSE	1.71	1.75	1.59	1.63	2.65	2.77
	ME	0.21	0.17	0.11	0.09	-0.55	-0.54
SEQ	MAE	1.33	1.33	1.45	1.26	1.60	1.90
	RMSE	1.73	1.79	1.97	1.80	2.23	2.54
Predictive	VAR vs. SEQ	-0.391	-0.098	-0.883	0.001	0.158	-0.458
accuracy	p-values	0.698	0.922	0.383	0.999	0.876	0.650
	VAR vs. SEQ	1.683	2.011	3.155	1.312	1.945	4.610
Encompassing	p-values	0.101	0.052	0.003	0.198	0.060	0.000
	SEQ vs. VAR	1.459	2.010	4.190	2.893	0.370	0.847
	p-values	0.154	0.052	0.000	0.007	0.714	0.403

Note. ME: Mean Error. MAE: Mean Absolute Error. RMSE: Root Mean Square Error.

$t-1$. An alternative strategy that bypasses this problem is based on building six single-equation forecasting models, one for each forecasting horizon, such that at every step the most comprehensive available information set is used. The single-equation model used to produce the k-step ahead forecast for $\Delta_{12}IPI_t$ can be written as

$$\begin{aligned} \Delta_{12}IPI_t &= a_k(L)\Delta_{12}IPI_{t-k} + b_{k,1}(L)\Delta_{12}TON_{t-k+1} + \\ &\quad + b_{k,2}(L)\Delta_{12}PP_{t-k+1} + \boldsymbol{\Upsilon}_k'\mathbf{d}_t + \upsilon_{k,t} \end{aligned} \tag{4}$$

where $\mathbf{d}_t$ is the same as in (1) when $k = 1, 2$. At longer horizons ($k = 3, ..., 6$) $\mathbf{d}_t$ does not include the dummies for August and December. In fact, the single-equation framework does not allow to produce forecasts for PP, so that August and December dummies cannot be produced either.

This procedure is related to the concept of "dynamic estimation", which sometimes can be useful for forecasting in the presence of model misspecifications (Clements and Hendry, 1998).

The equations are initially over-parameterized, and are successively recursively simplified in the usual general-to-specific framework. The final 6 equations selected to produce the 1 to 6-step ahead forecasts are reported in Appendix C. As might be easily expected, the forecasting equations for the longer horizons show residuals autocorrelation at lower lags.

The forecasting performance of the single-equation models, summarized in Table 3, is slightly worse than that of the VAR. Only the five- and six-step ahead RMSEs are slightly better than the corresponding VAR statistics.

Table 4: Forecasts evaluation statistics: VAR vs CSC and IRS.

Steps ahead		1	2			1	2
	ME	0.22	0.12		ME	0.13	0.14
VAR	MAE	1.23	1.21	VAR	MAE	1.29	1.23
	RMSE	1.71	1.67		RMSE	1.74	1.69
	ME	0.00	-0.64		ME	-1.06	-1.25
CSC	MAE	1.09	1.31	IRS	MAE	1.56	1.90
	RMSE	1.61	1.80		RMSE	2.07	2.34
Predictive	VAR vs. CSC	0.586	-0.353		VAR vs. IRS	-0.914	-1.952
accuracy	p-values	0.562	0.726		p-values	0.367	0.060
	VAR vs. CSC	2.933	3.100		VAR vs. IRS	2.978	2.442
Encompassing	p-values	0.006	0.004		p-values	0.005	0.021
	CSC vs. VAR	2.042	2.736		IRS vs. VAR	2.849	3.484
	p-values	0.049	0.010		p-values	0.007	0.001

Note. ME: Mean Error. MAE: Mean Absolute Error. RMSE: Root Mean Square Error. The statistics on the VAR are computed on the same samples covered by CSC and IRS, respectively, by excluding the months for which CSC or IRS did not publish their forecasts.

However, the MAE of the six-step ahead predictions is smaller for the VAR, consistently with the previous observation that the deterioration of the VAR forecasting statistics at the longest horizons is due to one large forecast error. Finally, the tests for predictive accuracy are never significant and those for encompassing do not show a clear pattern.

All in all, we believe that even if the VAR performance is not dramatically superior to the single-equations' one, nevertheless the VAR framework is neater and more flexible, so that we prefer again our VAR to the single-equation alternatives. Finally, from a very practical standpoint, the VAR has also the advantage of "protecting" the forecasting procedure from occasional delays in the availability of TON and PP.

3.3 CSC and IRS forecasts

The comparison with the CSC and IRS projections is limited to the two-step ahead forecast horizons, but given the high reputation of these two forecasters, it is for us very important to obtain a similar forecasting record on this short horizon. The computations for the MAE and the RMSE are carried out considering the same sample for the VAR and CSC and IRS, respectively. The results of the comparisons are reported in Table 4. The outcomes suggest that our VAR behaves slightly better than the CSC survey on the two-step ahead horizon, but is somewhat inferior on the one-step ahead. According to the predictive accuracy tests, these differences are not statistically significant, however. On the contrary, our predictions appear to be consistently superior to IRS', particularly so at the two-step ahead horizon, where the predictive accuracy test is marginally significant. The encompassing tests always reject.

4 Using the forecasts to improve the quality of trend-cycle estimates

The overall evidence collected in the previous section suggests that we can fruitfully use our VAR to forecast the Italian industrial production index. Now we want to show that our forecasts, which we believe are useful *per se*, can be used also in a different context to retrieve useful information and improve the analysis of the industrial business cycle.

The estimation of a cyclical indicator from an observed series can be typically represented as a filtering problem. As such, the indicator itself is calculated using a weighted average of past, present, and future values of the variable by applying a symmetric filter to the observed series. However, at the end of the series, say at time T, no future values are available, and asymmetric uni-lateral filters must be used to obtain preliminary estimates to be revised later. It should be highlighted that the most recent part of the cyclical indicator series is the most interesting to business analysts and policy makers. Therefore it is of paramount importance to minimize the revisions in the cyclical indicator so to obtain a series that is reliable and stable. In this context, the forecasts can be used to extend the observed series so as to apply a symmetric filter also at the end-points, obtaining a cyclical indicator less affected by revisions.

This is done in a spirit similar to the X-12-ARIMA seasonal adjustment procedure, where univariate forecasts are used to robustify the procedure, in order to get less revised seasonal factors (Findley *et al.*, 1998). Note that, again in analogy with X-12-ARIMA, there is no need that the forecasts and the indicators be derived within the same framework. For example, the selected filtering tool can be either parametric (as in this paper) or nonparametric (as it was in a previous draft). The filtering tool is not directly related to the forecasts, being specifically designed only to solve the filtering problem of interest, not the forecasting problem.

4.1 Characterization of the problem

An important feature of the industrial production index is its strong cyclical behavior. Indeed, the down- and upturns of the industrial production cycle are of major interest to business analysts. In order to highlight this feature, the unobserved component framework seems the most appropriate one. Let the series IPI be composed of four (unobserved) components:

$$IPI_t = \mu_t + \omega_t + \gamma_t + \varepsilon_t \qquad (5)$$

which correspond to the trend-cycle (μ_t), the seasonal (ω_t), the working days (γ_t), and the irregular component (ε_t), respectively.

A cyclical indicator can be built by eliminating the seasonal and the other short term movements from the series, leaving only the trend-cycle.[18] In order to estimate the latter, many methods have been proposed. Here we adopt the structural time series approach (see Harvey, 1989), which specifies a statistical model for each of the components in (5). The explicit specification of a statistical model for the components is often regarded as a key advantage of the structural time series approach over non parametric approaches and model based approaches based on reduced forms (*e.g.* the ARIMA model-based approach to decompose time series; see Maravall, 1995, and the references therein).[19]

4.2 The structural time series model

As suggested in the previous subsection, one advantage of the model-based approach is the explicit specification of the models for the components, making it possible to obtain components with the desired characteristics. Let us consider again the representation (5), and let ε_t be $\mathsf{IID}(0, \sigma^2_\varepsilon)$. Since we are particularly interested in obtaining a smooth trend-cycle component for easily detecting expansion and recession periods, we choose to model μ_t as an Integrated Random Walk (IRW) (Pedregal and Young, 2002):

$$\begin{aligned} \mu_t &= \mu_{t-1} + \beta_{t-1} \\ \beta_t &= \beta_{t-1} + \zeta_t \end{aligned} \qquad (6)$$

where ζ_t is $\mathsf{IID}(0, \sigma^2_\zeta)$. The variable β_t is generally called the "slope" or the "trend derivative". Even if the specification of the trend as in (6) implies

[18] In this paper we are not interested in disentangling long term trend and business cycle oscillations, given that the data we are dealing with feature frequent cycles in the classical sense, so that detrending is not necessary. Indeed, whenever we refer to business cycle, we do so in the sense of *classical cycle* and not in the sense of *growth cycle* (which considers detrended series). At any rate, 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.

[19] Indeed, in an early version of this paper (documenti di lavoro ISAE, 20/01) we used the ARIMA model-based approach, obtaining qualitatively similar results.

an $\mathsf{I}(2)$ representation for IPI_t, nevertheless it represents a useful and widespread tool in the context of trend extraction problems, because it produces a very smooth trend component. In fact, the problem of dating expansion and recession in monthly time series can be a very difficult task and the IRW trend, being much smoother than other alternatives, allows to accomplish this task in an easy and effective way, simply using the trend derivative definition (García-Ferrer and Bujosa-Brun, 2000).[20] In particular, period t is defined as an expansion (recession) if $\mu_t - \mu_{t-1} = \beta_{t-1} > 0$ (< 0).

Concerning the other components, the seasonal pattern is defined in trigonometric form:

$$\omega_t = \sum_{j=1}^{6} \omega_{jt} \tag{7}$$

with

$$\begin{aligned} \omega_{jt} &= \omega_{j,t-1}\cos\lambda_j + \omega^*_{j,t-1}\sin\lambda_j + \kappa_{jt} \\ \omega^*_{jt} &= -\omega_{j,t-1}\sin\lambda_j + \omega^*_{j,t-1}\cos\lambda_j + \kappa^*_{jt} \end{aligned} \tag{8}$$

where κ_{jt} and κ^*_{jt} $\mathsf{IID}(0, \sigma^2_{j\kappa})$. The parameter $\lambda_j = 2\pi j/12$ represents the frequency in radians.

Finally, the working days effect is modelled as:

$$\gamma_t = \delta TD_t \tag{9}$$

where TD_t is again the number of working days in month t and δ is a coefficient to be estimated. A time varying structure for δ has also been tested but it did not prove necessary.

The error terms ε_t, ζ_t, κ_{jt} and κ^*_{jt} are mutually independent at all leads and lags. The model composed by equations (5) to (9) can be easily put in state space form and the further assumption of normality allows estimation by maximum likelihood of the unknown variances and of δ (so called *hyper-parameters*) by means of the Kalman filter. The Kalman smoother can then be used to estimate the different components.[21]

[20] Pedregal and Young (2002) note that the IRW trend has the properties of a cubic spline. We have also tested the alternative $\mathsf{I}(1)$ specification with $\beta_t = \phi\beta_{t-1} + \zeta_t$, and $|\phi| < 1$. Though the results are qualitatively similar to those illustrated here, nevertheless the resulting trend is rougher, rendering more difficult the detection of the business cycle phases.

[21] For the practical implementation we use Ox 3.0 (Doornik, 2001) and the routine SsfPack 2.2 (Koopman *et al.*, 1998).

Table 5: Structural model estimates

$\sigma^2_\zeta = 0.0152$	$\sigma^2_{\kappa_1} = 0.0037$	$\sigma^2_{\kappa_2} = 0.0019$	$\sigma^2_{\kappa_3} = 0.0052$
$\sigma^2_{\kappa_4} = 0.0072$	$\sigma^2_{\kappa_5} = 0.0014$	$\sigma^2_{\kappa_6} = 0.0017$	$\delta = \underset{(0.10885)}{2.7128}$
Ljung-Box(24)	0.5234		
Normality	0.1118		

The variances are expressed in terms of signal-to-noise ratios (*i.e.* are all normalized by σ^2_ε). δ is the coefficient attached to the trading days term; the number in brackets is the standard error. The last two rows report the p-values of the Ljung-Box and normality tests on residuals.

4.3 Reducing revisions

Classical formulae for signal extraction allow us to recover the minimum mean square linear estimator of the unobserved components β_t and μ_t, given an infinite realization of the observed series y_t:

$$\widehat{\beta}_{t|\infty} = \sum_{j=-\infty}^{\infty} w_j y_{t+j} \tag{10}$$

$$\widehat{\mu}_{t|\infty} = \sum_{j=-\infty}^{\infty} u_j y_{t+j} \tag{11}$$

where the w's and the u's are two set of weights. In practice, we never observe infinite samples; instead, in the presence of a sample of finite length $T \geq t$ we have a sequence of preliminary estimates $\widehat{\beta}_{t|t+k}$ and $\widehat{\mu}_{t|t+k}$ $(0 \leq k \leq T-t)$. When $k = 0$ we have the so called *concurrent estimate*; adding observations after time t produces a set of estimates which converge to their "final" value as t grows in such a way that $t + k \to T \gg t$.

In this section we show that the use of the forecasts coming out from the model described in Section 2 dramatically improve the reliability of the preliminary estimates of the trend and of the slope components of IPI, making them much better devices in order to monitor the business cycle in the manufacturing sector. Indeed, Bruno (2001) shows that the revisions of conventional estimates of the trend component of IPI may be unacceptably large for the purpose of analysing the current industrial business cycle situation.

In order to check the importance of revisions in the trend estimates, and to evaluate the advantages deriving from using our model's forecasts, we perform a historical simulation from September 1995 to December 2001, estimating the trend and slope components at every period, for the original series and for the one extended, respectively, with six-step ahead ARIMA forecasts and with six-step ahead VAR forecasts.[22]

The quantities

$$\begin{aligned} r_{\mu,k} &= \widehat{\mu}_{t|t+k} - \widehat{\mu}_{t|t+k-1} \qquad k = 1, 2, ... &(12)\\ r_{\beta,k} &= \widehat{\beta}_{t|t+k} - \widehat{\beta}_{t|t+k-1} &(13)\end{aligned}$$

represent, for every k, the monthly revisions in $\widehat{\mu}_t$ and $\widehat{\beta}_t$, k months after the concurrent estimate.

We compute (12) and (13) for every month from October 1995 onward, obtaining a distribution of revisions for every k ranging from 1 (with 75 observations) to 75 (just one observation). In practice we are usually interested in small values of k, say $k \leq 10$. 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 forecasts improves on the revision process. Figures 2 and 3 show clearly how effective is the improvement on the revision pattern using the forecasts from our model. The black bar (labelled "No forecasts") is the variance of the monthly revisions after k periods (x-axis) using the standard procedure that relies only on historical data. Note that converge of revisions is faster for the trend than for the slope component. In both cases the extension of the original series with the ARIMA forecasts does not improve significantly the revision pattern. On the contrary, revisions diminish dramatically using the VAR forecasts; trend revision variance is reduced by over 50% during the first three periods and by 40% during the fourth. As far as slope revisions are considered, the reduction is greater than 50% during the first seven periods.

4.4 Detection of turning points

In order to further assess how important is the improvement in the revision process showed in the previous sub-section, it is possible to analyze if it helps in detecting turning points earlier. In order to do this we use the results of the simulation of the previous sub-section together with the rule that a contraction (expansion) takes place in period t as long as $\beta_{t-1} < 0$ (> 0), as suggested by García-Ferrer and Bujosa-Brun (2000). As a reference, we consider the turning points identified with the observed series ending in April 2002 as "final".

It is important to stress that our aim is not to find out the best approach to signal turning points, nor to forecast them. In fact, turning points are often

[22]The ARIMA model is the same as the one described in section 3.

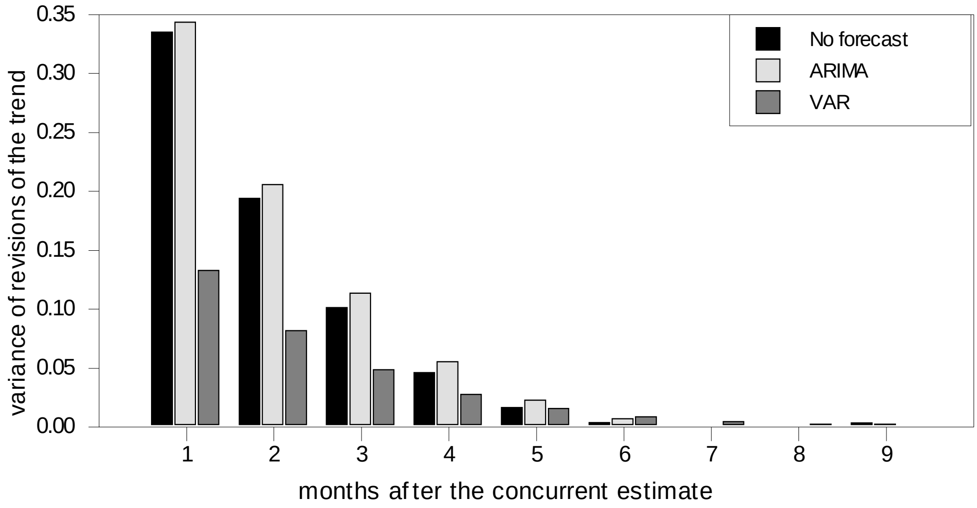

Figure 2: Variance of revisions of the trend (μ_t)

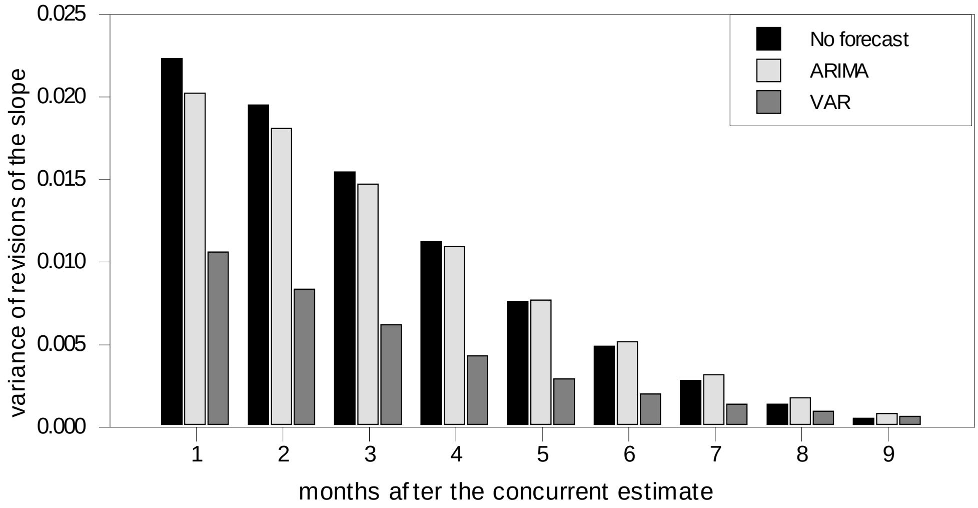

Figure 3: Variance of revisions of the slope (β_t)

Table 6: Detection dates of turning points with different forecasts

Turning points locations	No forecasts	ARIMA	VAR
1995:10 (p)	6	7	2
1996:9 (t)	7	7	4
1998:2 (p)	6	3	2
1999:2 (t)	6	1	2
2000:11 (p)	5	5	2
Mean lag	6.0	4.6	2.4

The table reports the delay, in months, 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 April 2002. Mean lag is the average lag of the detection with different methods. Six forecasts added to time series for ARIMA and VAR.

claimed to be essentially non-linear phenomena, and non-linear methods seem the most natural candidates to forecast them (Camacho and Perez-Quiros, 2002). On the contrary, we want to show that our forecasts can be utilized to recognize turning points earlier, using a standard procedure to detect them.[23] This is also to say that at this stage we are not particularly concerned with the determination of the exact or "true" timing of the turning points. Rather, we want to show that the results of a standard routine to identify turning points can be significantly improved by embodying our forecasts.

The turning points identified by the procedure over the period 1995.9-2001.12 using the actual data up to April 2002 are five (three peaks and two troughs): they are not so many, but going back further would have led to a too pronounced loss of degrees of freedom for our model.[24] Table 6 shows the main results: the dates in the first column represent the turning points estimated as of April 2002.

The mean lag in the detection of turning points with the original series is 6 months. The use of the ARIMA model forecasts in some cases reduces the time needed to detect turning points: the reduction, on average, is 1.4 months. A further improvement is obtained with the VAR forecasts,[25] which

[23] In a previous version of this paper turning points were identified applying the procedure proposed by Bry and Boschan (1971) to the trend-cycle series extracted using TRAMO-SEATS (Gómez & Maravall, 1998). The dates of the turning points identified using that procedure are very similar to those reported here.

[24] By replicating a historical situation, we estimate the model using information up to time t.

[25] Given that we have only five observations, the calculation of the mean could be ques-

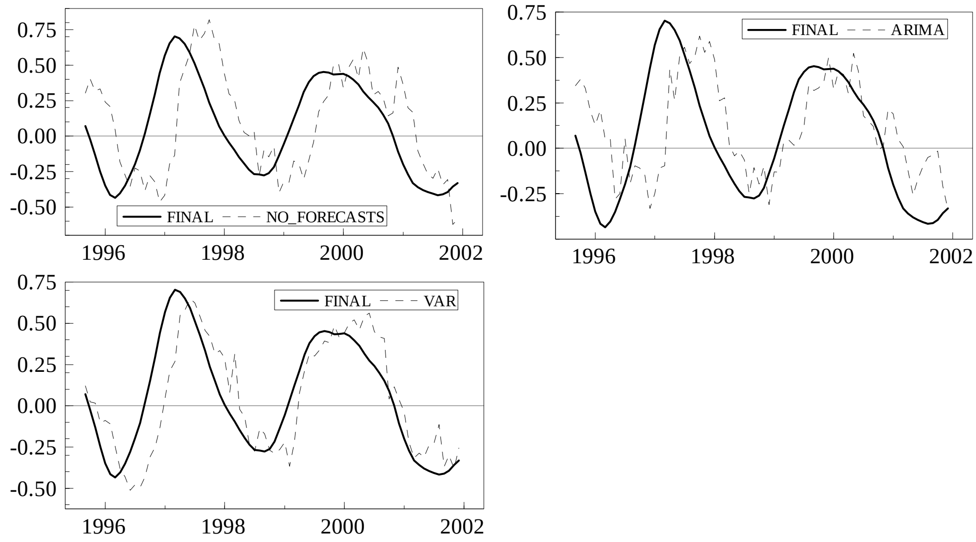

Figure 4: Concurrent estimate of the slope $\beta_{t|t}$ obtained with the original data ('No forecasts'), with the forecasts produced with the ARIMA model and with those stemming from the VAR model. The final slope ('Final') $\beta_{t|T}$ is estimated using the entire sample and is reported for comparison.

in four out of five cases allow us to detect a turning point after only two months, while in one case the time elapsed is four months.

Figure 4 shows the concurrent estimate of the slope for every t ranging from 1995.9 to 2001.12. In the upper left panel the conventional estimate of the slope is shown; in the upper right panel is displayed the concurrent estimate of the slope obtained by appending to the original data the ARIMA forecasts. Finally, the lower panel shows the concurrent estimate of the slope obtained by adding the VAR forecasts. In each case the "final" estimate of the slope $\beta_{t|T}$ (that is the one estimated using the whole sample ending in April 2002) is reported for comparison. It is clear that the conventional slope ("No forecasts") performs worse than the others in capturing turning points. On the other hand, the slope obtained using ARIMA forecasts does not show a consistent advantage over the "No forecast" case, also considering that in some cases it gives false signals. In the end, the slope derived using the VAR forecasts performs consistently better than the other two over the sample examined, being much closer to the "final" one, and not giving false signals.

tioned. However, it should be stressed that we find a monotonic decrease in the delay with which turning points are detected using our forecasts for all the turning points considered here. This is well reflected by the mean lag.

5 Concluding remarks

In this paper we propose a simple model to forecast the 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 and of simple single-equation alternatives: in addition, they 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 our model we are able to produce reliable forecasts on longer horizons, which is one of our main goals.

We argue that our forecasts, which are useful *per se*, can also be used to improve significantly on the quality and reliability of the estimates of a cyclical indicator obtained using signal-extraction (smoothing) techniques. In particular, we compare the variance of the revisions of a cyclical indicator estimated using our model'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 VAR forecasts is 3.6 months, that is a 60% reduction. We guess that a clear indication to practitioners 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

[1] 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. (requests may be addressed to `relazioni.esterne@isae.it`)

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

[3] Bierens, H.J. & Guo, S. (1993): Testing stationarity and trend stationarity against the unit root hypothesis, *Econometric Reviews*, 12,1-32.

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

[5] Bruno, G. (2001). Seasonal adjustment of Italian industrial production index using TRAMO-SEATS, ISAE, Documenti di lavoro n. 18/01. (requests may be addressed to `relazioni.esterne@isae.it`)

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

[7] Camacho, M. & Perez-Quiros, G. (2002). This is what the leading indicators lead. *Journal of Applied Econometrics*, 17, 61-80.

[8] Cecchetti, S.G., Kashyap, A.K. & Wilcox, D.W. (1997). Interactions between the Seasonal and Business Cycles in Production and Inventories. *American Economic Review*, 87, 884-892.

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

[10] Clements, M.P. & Hendry, D.F. (1998). *Forecasting Economic Time Series*, Cambridge: Cambridge University Press.

[11] Clements, M.P. & Hendry, D.F. (1999). *Forecasting Non-stationary Economic Time Series*, Cambridge (Mass.): MIT Press.

[12] 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.

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

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

[15] Doornik, J.A. (2001). *Ox 3.0 - An Object-Oriented Matrix Programming language*. Timberlake Consultants Press.

[16] Doornik, J.A. & Hendry, D.F. (2001). *Modelling Dynamic Systems Using PcGive, Vol II,* London: Timberlake Consultants Ltd.

[17] 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.

[18] 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. (`http://www.eur.nl/WebDOC/doc/econometrie/feweco20010426095757.pdf`)

[19] García-Ferrer, A. & Bujosa-Brun, M. (2000). Forecasting OECD industrial turning points using unobserved components models with business survey data, *International Journal of Forecasting*, 16, 207-227.

[20] García-Ferrer, A. & Queralt, R.A. (1997). A note on forecasting international tourism demand in Spain, *International Journal of Forecasting*, 13, 539-549.

[21] Gómez, V., & Maravall, A. (1998). *Programs TRAMO and SEATS: Instructions for the User,* Madrid: Banco de España. (`http://www.bde.es/servicio/software/econome.htm`)

[22] 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. (`http://www.fief.se/library/wp/WP150rev.pdf`)

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

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

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

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

[27] Koopman, S.J., Shephard, N. & Doornik J.A. (1998). Statistical algorithms for models in state space using SsfPack 2.2. *Econometrics Journal*, n. 1, 1-55.

[28] Krolzig, H.M. (2001). General-to-specific reductions of Vector Autoregressive Processes. In Friedmann, R., L. Knüppel & H. Lütkepohl (eds.), *Econometric Studies - A Festschrift in Honour of Joachim Frohn*, pp. 129-157, Münster, LIT Verlag.

[29] 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. (`http://swopec.hhs.se/hastef/papers/hastef0418.pdf`).

[30] Maravall, A. (1995). Unobserved components in economic time series. In H. Pesaran and M.M. Wickens (Eds.), *Handbook of Applied Econometrics*, Vol. 1. Oxford: Basil Blackwell.

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

[32] McGuckin, R.H., Ozyildirim, A., & Zarnowitz, V. (2001). The composite index of leading economic indicators: how to make it more timely. NBER Working Paper 8430.

[33] Osborn, D.R. (2002). Unit-root versus deterministic representations of seasonality for forecasting. In M.P. Clements and D.F. Hendry (Eds.), *A Companion to Economic Forecasting*. Oxford: Basil Blackwell.

[34] Osborn, D.R., Heravi, S. & Birchenhall, C.R. (1999). Seasonal unit roots and forecasts of two-digit European industrial production, *International Journal of Forecasting*, 15, 27-47.

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

[36] Pedregal, D.J. & Young, P.C. (2002). Statistical Approaches to Modeling and Forecasting Time Series. In M.P. Clements and D.F. Hendry (Eds.), *A Companion to Economic Forecasting*. Oxford: Basil Blackwell.

[37] 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.

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

[39] Simpson, P.W., Osborn, D.R. & Sensier, M. (2001). Forecasting UK Industrial Production over the Business Cycle, *Journal of Forecasting*, 20, 405-424.

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

[41] U.S. Census Bureau (2002). X-12-ARIMA Reference Manual. (`http://www.census.gov/srd/www/x12a/`).

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

A Unit roots

Despite the practical approach adopted in the main text, in this appendix we report for completeness the results of conventional tests for unit roots in seasonal time series. However, the tests that follows are valid under the maintained hypothesis of no structural breaks and pure AR representations. This conflicts somewhat with our expectations.

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 (Beaulieu and Miron, 1993). The results reported in Table 7 indicate that the presence of all the twelve unit roots is strongly rejected for all the series, but the presence of unit roots at least at some frequencies can never be excluded.

In order to investigate the presence of possible multiple unit roots at the zero frequency, we carry out formal tests on the annual differences of the series. The results of these tests are reported in Table 8. Both the unit roots and the stationarity tests seem to indicate that the annual differences of our series do not contain unit roots at frequency zero.

Table 7: 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.

Table 8: Tests for unit roots in seasonally differenced series

	ADF_c		$\text{ADF}_{c,t}$		BG2	BG3	BG4	BG5	BG6
$\Delta_{12}IPI$	-3.589	**	-3.566	*	0.061	0.061	0.061	0.860	0.713
$\Delta_{12}TON$	-4.105	**	-4.332	**	0.318	0.315	0.316	0.791	0.114
$\Delta_{12}PP$	-3.684	**	-3.666	*	0.632	0.586	0.588	0.328	0.219

Δ_{12} denotes the seasonal difference. The table reports the results from the Augmented Dickey-Fuller (with constant, ADF_c, and with constant and trend, $\text{ADF}_{c,t}$) tests for zero-frequency unit roots and from the Bierens-Guo (BG2-BG6) tests for stationarity (Bierens and Guo, 1993) applied to the seasonally differenced series. For the ADF tests, "*" and "**" denote values that are significant at 5% and 1%, respectively. For the Bierens-Guo tests, p-values are reported.

B VAR estimates

For completeness, in this Appendix we report the estimated VAR. The VAR diagnostics are reported in Section 2. The estimation period is 1988:03-1998:12. 36 out of sample observations (1999:01-2001:12) have been left for forecast evaluation.

$$
\begin{aligned}
\Delta\Delta_{12} IPI = & - \underset{(0.104)}{0.224}\ \Delta\Delta_{12} IPI_{t-1} - \underset{(0.0967)}{0.194}\ \Delta\Delta_{12} IPI_{t-2} - \underset{(0.0926)}{0.182}\ \Delta\Delta_{12} IPI_{t-3} \\
& - \underset{(0.0832)}{0.0153}\ \Delta\Delta_{12} IPI_{t-4} - \underset{(0.0611)}{0.00757}\ \Delta\Delta_{12} IPI_{t-5} + \underset{(0.0484)}{0.114}\ \Delta\Delta_{12} IPI_{t-9} \\
& - \underset{(0.0555)}{0.00307}\ \Delta\Delta_{12} IPI_{t-12} - \underset{(0.0581)}{0.0803}\ \Delta\Delta_{12} IPI_{t-13} - \underset{(0.0628)}{0.0971}\ \Delta\Delta_{12} TON_{t-1} \\
& - \underset{(0.0601)}{0.0964}\ \Delta\Delta_{12} TON_{t-2} - \underset{(0.0557)}{0.0321}\ \Delta\Delta_{12} TON_{t-3} - \underset{(0.0482)}{0.103}\ \Delta\Delta_{12} TON_{t-4} \\
& - \underset{(0.0401)}{0.0474}\ \Delta\Delta_{12} TON_{t-5} - \underset{(0.0328)}{0.0891}\ \Delta\Delta_{12} TON_{t-9} + \underset{(0.0369)}{0.0718}\ \Delta\Delta_{12} TON_{t-12} \\
& + \underset{(0.0399)}{0.0666}\ \Delta\Delta_{12} TON_{t-13} - \underset{(0.0235)}{0.0297}\ \Delta\Delta_{12} PP_{t-1} - \underset{(0.0225)}{0.000546}\ \Delta\Delta_{12} PP_{t-2} \\
& + \underset{(0.0214)}{0.0227}\ \Delta\Delta_{12} PP_{t-3} + \underset{(0.021)}{0.0248}\ \Delta\Delta_{12} PP_{t-4} - \underset{(0.0179)}{0.0105}\ \Delta\Delta_{12} PP_{t-5} \\
& + \underset{(0.0156)}{0.023}\ \Delta\Delta_{12} PP_{t-9} - \underset{(0.0172)}{0.0116}\ \Delta\Delta_{12} PP_{t-12} - \underset{(0.0174)}{0.0179}\ \Delta\Delta_{12} PP_{t-13} \\
& - \underset{(0.0956)}{0.485}\ \Delta_{12} IPI_{t-1} + \underset{(0.0579)}{0.115}\ \Delta_{12} TON_{t-1} + \underset{(0.0158)}{0.0333}\ \Delta_{12} PP_{t-1} \\
& + \underset{(0.00297)}{0.00146} - \underset{(0.0248)}{0.0906}\ d918_t + \underset{(0.0254)}{0.0392}\ d958_t \\
& - \underset{(0.0231)}{0.00564}\ d9211_t - \underset{(0.0118)}{0.0388}\ dumaugn_t + \underset{(0.00999)}{0.0549}\ dumaugp_t \\
& - \underset{(0.00947)}{0.0219}\ dumdecn_t + \underset{(0.0106)}{0.0209}\ dumdecp_t + \underset{(0.0541)}{0.474}\ \Delta_{12} \log TD_t \\
& - \underset{(0.0745)}{0.271}\ \Delta_{12} \log TD_{t-1}
\end{aligned}
$$

$$
\begin{aligned}
\Delta\Delta_{12}TON \quad = \quad & \underset{(0.226)}{0.392}\ \Delta\Delta_{12}IPI_{t-1} + \underset{(0.211)}{0.0571}\ \Delta\Delta_{12}IPI_{t-2} + \underset{(0.202)}{0.0391}\ \Delta\Delta_{12}IPI_{t-3} \\
& + \underset{(0.181)}{0.0613}\ \Delta\Delta_{12}IPI_{t-4} + \underset{(0.133)}{0.0797}\ \Delta\Delta_{12}IPI_{t-5} + \underset{(0.105)}{0.0558}\ \Delta\Delta_{12}IPI_{t-9} \\
& + \underset{(0.121)}{0.0168}\ \Delta\Delta_{12}IPI_{t-12} - \underset{(0.127)}{0.0234}\ \Delta\Delta_{12}IPI_{t-13} - \underset{(0.137)}{0.453}\ \Delta\Delta_{12}TON_{t-1} \\
& - \underset{(0.131)}{0.156}\ \Delta\Delta_{12}TON_{t-2} - \underset{(0.121)}{0.0877}\ \Delta\Delta_{12}TON_{t-3} - \underset{(0.105)}{0.113}\ \Delta\Delta_{12}TON_{t-4} \\
& - \underset{(0.0875)}{0.187}\ \Delta\Delta_{12}TON_{t-5} - \underset{(0.0716)}{0.107}\ \Delta\Delta_{12}TON_{t-9} - \underset{(0.0804)}{0.272}\ \Delta\Delta_{12}TON_{t-12} \\
& - \underset{(0.0869)}{0.0857}\ \Delta\Delta_{12}TON_{t-13} - \underset{(0.0512)}{0.0874}\ \Delta\Delta_{12}PP_{t-1} + \underset{(0.0491)}{0.0391}\ \Delta\Delta_{12}PP_{t-2} \\
& - \underset{(0.0467)}{0.00252}\ \Delta\Delta_{12}PP_{t-3} + \underset{(0.0458)}{0.00439}\ \Delta\Delta_{12}PP_{t-4} + \underset{(0.039)}{0.00667}\ \Delta\Delta_{12}PP_{t-5} \\
& + \underset{(0.0339)}{0.0107}\ \Delta\Delta_{12}PP_{t-9} - \underset{(0.0376)}{0.027}\ \Delta\Delta_{12}PP_{t-12} - \underset{(0.038)}{0.0484}\ \Delta\Delta_{12}PP_{t-13} \\
& - \underset{(0.208)}{0.361}\ \Delta_{12}IPI_{t-1} - \underset{(0.126)}{0.22}\ \Delta_{12}TON_{t-1} + \underset{(0.0345)}{0.0957}\ \Delta_{12}PP_{t-1} \\
& + \underset{(0.00646)}{0.015} - \underset{(0.054)}{0.0309}\ d918_t + \underset{(0.0553)}{0.0415}\ d958_t \\
& - \underset{(0.0503)}{0.0913}\ d9211_t - \underset{(0.0257)}{0.0463}\ dumaugn_t + \underset{(0.0218)}{0.0206}\ dumaugp_t \\
& - \underset{(0.0206)}{0.0568}\ dumdecn_t + \underset{(0.023)}{0.0337}\ dumdecp_t + \underset{(0.118)}{0.354}\ \Delta_{12}\log TD_t \\
& - \underset{(0.162)}{0.501}\ \Delta_{12}\log TD_{t-1}
\end{aligned}
$$

$$
\begin{aligned}
\Delta\Delta_{12}PP \;=\;\; & \underset{(0.541)}{0.482}\ \Delta\Delta_{12}IPI_{t-1} + \underset{(0.504)}{0.729}\ \Delta\Delta_{12}IPI_{t-2} + \underset{(0.483)}{0.678}\ \Delta\Delta_{12}IPI_{t-3} \\
& + \underset{(0.434)}{0.358}\ \Delta\Delta_{12}IPI_{t-4} + \underset{(0.318)}{0.000218}\ \Delta\Delta_{12}IPI_{t-5} - \underset{(0.252)}{0.19}\ \Delta\Delta_{12}IPI_{t-9} \\
& - \underset{(0.289)}{0.164}\ \Delta\Delta_{12}IPI_{t-12} - \underset{(0.303)}{0.108}\ \Delta\Delta_{12}IPI_{t-13} + \underset{(0.327)}{0.486}\ \Delta\Delta_{12}TON_{t-1} \\
& - \underset{(0.313)}{0.052}\ \Delta\Delta_{12}TON_{t-2} - \underset{(0.29)}{0.168}\ \Delta\Delta_{12}TON_{t-3} - \underset{(0.251)}{0.118}\ \Delta\Delta_{12}TON_{t-4} \\
& - \underset{(0.209)}{0.132}\ \Delta\Delta_{12}TON_{t-5} - \underset{(0.171)}{0.0539}\ \Delta\Delta_{12}TON_{t-9} + \underset{(0.192)}{0.171}\ \Delta\Delta_{12}TON_{t-12} \\
& + \underset{(0.208)}{0.0763}\ \Delta\Delta_{12}TON_{t-13} - \underset{(0.122)}{0.227}\ \Delta\Delta_{12}PP_{t-1} - \underset{(0.118)}{0.124}\ \Delta\Delta_{12}PP_{t-2} \\
& - \underset{(0.112)}{0.0214}\ \Delta\Delta_{12}PP_{t-3} + \underset{(0.11)}{0.0651}\ \Delta\Delta_{12}PP_{t-4} + \underset{(0.0932)}{0.297}\ \Delta\Delta_{12}PP_{t-5} \\
& - \underset{(0.0812)}{0.0602}\ \Delta\Delta_{12}PP_{t-9} - \underset{(0.0899)}{0.354}\ \Delta\Delta_{12}PP_{t-12} - \underset{(0.0909)}{0.0291}\ \Delta\Delta_{12}PP_{t-13} \\
& - \underset{(0.499)}{0.68}\ \Delta_{12}IPI_{t-1} - \underset{(0.302)}{0.17}\ \Delta_{12}TON_{t-1} + \underset{(0.0825)}{0.025}\ \Delta_{12}PP_{t-1} \\
& + \underset{(0.0155)}{0.0225} - \underset{(0.129)}{0.062}\ d918_t + \underset{(0.132)}{0.0209}\ d958_t \\
& - \underset{(0.12)}{0.409}\ d9211_t - \underset{(0.0616)}{0.0382}\ dumaugn_t + \underset{(0.0521)}{0.00258}\ dumaugp_t \\
& + \underset{(0.0494)}{0.106}\ dumdecn_t - \underset{(0.0551)}{0.108}\ dumdecp_t - \underset{(0.282)}{0.836}\ \Delta_{12}\log TD_t \\
& - \underset{(0.388)}{0.354}\ \Delta_{12}\log TD_{t-1}
\end{aligned}
$$

C Single-equation models

C.1 One-step ahead

$$\begin{aligned}
\Delta_{12}IPI \;=\;& \underset{(0.0623)}{0.255}\ \Delta_{12}IPI_{t-1} - \underset{(0.00251)}{0.00192} + \underset{(0.0332)}{0.165}\ \Delta_{12}TON_t \\
&+ \underset{(0.0253)}{0.15}\ \Delta_{12}TON_{t-12} + \underset{(0.0107)}{0.0538}\ \Delta_{12}PP_{t-3} + \underset{(0.0434)}{0.464}\ \Delta_{12}\log TD_t \\
&- \underset{(0.0561)}{0.193}\ \Delta_{12}\log TD_{t-1} - \underset{(0.0224)}{0.12}\ d918_t - \underset{(0.0107)}{0.0251}\ dumaugn_t \\
&+ \underset{(0.00873)}{0.0548}\ dumaugp_t
\end{aligned}$$

Test		Statistic	p-value
AR 1-12 test:	F(12,108)=	1.0617	0.3996
ARCH 1-12 test:	F(12,96) =	0.93271	0.5181
Normality test:	$\chi^2(2)$ =	1.3339	0.5133
hetero test:	F(15,104)=	1.1452	0.3269
hetero-X test:	F(38,81) =	1.5247	0.0573
RESET test:	F(1,119) =	1.6178	0.2059

C.2 Two-step ahead

$$\begin{aligned}
\Delta_{12}IPI \;=\;& \underset{(0.0464)}{0.112}\ \Delta_{12}IPI_{t-3} - \underset{(0.00253)}{0.00127} + \underset{(0.0352)}{0.119}\ \Delta_{12}TON_{t-1} \\
&+ \underset{(0.0273)}{0.0658}\ \Delta_{12}TON_{t-7} + \underset{(0.0285)}{0.14}\ \Delta_{12}TON_{t-12} + \underset{(0.0111)}{0.0801}\ \Delta_{12}PP_{t-2} \\
&+ \underset{(0.00929)}{0.0549}\ dumaugp_t - \underset{(0.024)}{0.127}\ d918_t + \underset{(0.0448)}{0.462}\ \Delta_{12}\log(TD)_t \\
&- \underset{(0.0479)}{0.196}\ \Delta_{12}\log(TD)_{t-1}
\end{aligned}$$

Test		Statistic	p-value
AR 1-12 test:	F(12,108)=	1.5245	0.1263
ARCH 1-12 test:	F(12,96) =	1.585	0.1088
Normality test:	$\chi^2(2)$ =	0.35396	0.8378
hetero test:	F(16,103)=	0.31639	0.9942
hetero-X test:	F(42,77) =	0.55803	0.9795
RESET test:	F(1,119) =	3.7923	0.0538

C.3 Three-step ahead

$$\begin{aligned}
\Delta_{12}IPI \;=\;& \underset{(0.00278)}{0.00418} + \underset{(0.0311)}{0.1}\ \Delta_{12}TON_{t-7} + \underset{(0.0325)}{0.175}\ \Delta_{12}TON_{t-12} \\
&+ \underset{(0.00953)}{0.122}\ \Delta_{12}PP_{t-2} + \underset{(0.0533)}{0.505}\ \Delta_{12}\log TD_t - \underset{(0.053)}{0.179}\ \Delta_{12}\log TD_{t-1}
\end{aligned}$$

Test		Statistic	p-value
AR 1-12 test:	F(12,112)=	1.9693	0.0335
ARCH 1-12 test:	F(12,100)=	2.6763	0.0037
Normality test:	$\chi^2(2)$ =	1.4538	0.4834
hetero test:	F(10,113)=	1.0161	0.4345
hetero-X test:	F(20,103)=	0.79015	0.7197
RESET test:	F(1,123) =	2.8985	0.0912

C.4 Four-step ahead

$$\begin{aligned}\Delta_{12}IPI = & \underset{(0.00306)}{0.0049} + \underset{(0.0371)}{0.101}\ \Delta_{12}TON_{t-3} + \underset{(0.03)}{0.163}\ \Delta_{12}TON_{t-12} \\ & + \underset{(0.0119)}{0.105}\ \Delta_{12}PP_{t-3} + \underset{(0.0559)}{0.519}\ \Delta_{12}\log TD_t - \underset{(0.0564)}{0.123}\ \Delta_{12}\log TD_{t-1}\end{aligned}$$

Test		Statistic	p-value
AR 1-12 test:	F(12,112)=	1.7406	0.0673
ARCH 1-12 test:	F(12,100)=	1.5212	0.1289
Normality test:	$\chi^2(2)$ =	8.5356	0.0140
hetero test:	F(10,113)=	1.1403	0.3390
hetero-X test:	F(20,103)=	0.85881	0.6378
RESET test:	F(1,123) =	0.60339	0.4388

C.5 Five-step ahead

$$\begin{aligned}\Delta_{12}IPI = & \underset{(0.049)}{0.18}\ \Delta_{12}IPI_{t-9} + \underset{(0.00271)}{0.0115} + \underset{(0.00934)}{0.115}\ \Delta_{12}PP_{t-4} \\ & + \underset{(0.0582)}{0.458}\ \Delta_{12}\log TD_t - \underset{(0.0574)}{0.166}\ \Delta_{12}\log TD_{t-1} + \underset{(0.0287)}{0.129}\ d958_t\end{aligned}$$

Test		Statistic	p-value
AR 1-12 test:	F(12,112)=	5.3293	0.0000
ARCH 1-12 test:	F(12,100)=	1.1673	0.3169
Normality test:	$\chi^2(2)$ =	13.794	0.0010
hetero test:	F(9,114) =	0.77248	0.6420
hetero-X test:	F(15,108)=	0.99552	0.4653
RESET test:	F(1,123) =	0.18651	0.6666

C.6 Six-step ahead

$$\begin{aligned}\Delta_{12}IPI = & \underset{(0.00269)}{0.0157} + \underset{(0.01)}{0.116}\ \Delta_{12}PP_{t-5} + \underset{(0.056)}{0.559}\ \Delta_{12}\log TD_t \\ & + \underset{(0.0309)}{0.164}\ I1988:8_t\end{aligned}$$

AR 1-12 test:	F(12,114)=	2.3406	0.0101
ARCH 1-12 test:	F(12,102)=	1.7139	0.0744
Normality test:	$\chi^2(2)$ =	8.7349	0.0127
hetero test:	F(5,120) =	0.20817	0.9585
hetero-X test:	F(6,119) =	0.34412	0.9121
RESET test:	F(1,125) =	0.18389	0.6688