A Robust and Forward-looking Industrial Production Indicator

MANMOHAN S SODHI, JASMINE SHARMA, SUKHMEET SINGH, ARVINDER WALIA

Against the backdrop of growing criticism of the index of industrial production, which provides information only about the past and sometimes fluctuates wildly, this article seeks to provide a more robust and forward-looking economic indicator of industrial growth. Such an indicator, based on past IIP numbers, can also serve as a benchmark for future IIP numbers when they are released. Using data on the IIP's three sub-series – manufacturing, mining, and electricity – it seeks to isolate the "noise" from the "signal" in two steps, enabling predictions for the two past months and four months into the future using the latest available IIP numbers in any given month.

1 Introduction

This study explores the creation of a forward-looking economic indicator to anticipate and understand industrial growth. Given the slow growth in India in 2013 relative to recent years, the importance of monitoring industrial growth has only increased. The question is how to do it. The macroeconomic indicators for monitoring the manufacturing sector (in particular) are gross domestic product (GDP), the purchasing managers index (PMI), and the index of industrial production (IIP). But GDP numbers appear only annually and the PMI is essentially an expectations-based one. Therefore, the IIP, which appears every month, albeit in arrears by two months, has become the most widely-used macroeconomic variable to monitor growth in industrial output. It comprises data from three major sectors weighted to give a composite number – manufacturing weighted (approximately) 75.5%, mining and quarrying 14.3%, and electricity 10.3%. Policymakers, statisticians, economists, analysts, planners, and business entities await and value the numbers as indicative of growth in industrial output and make decisions and pronouncements accordingly.[1]

However, there have been doubts over the IIP's relevance from a policy viewpoint. First, there is a substantial time lag involved. That is, the preliminary numbers are available with a lag of six weeks from the reference month, making these IIP numbers "retrospective". Second, there have been serious problems with the index as regards "noise" and month-on-month volatility in recent years, posing a credibility challenge in its use for monitoring growth and performance. For instance, in response to the revision of the January 2012 numbers in April, lowering year-on-year (y-o-y) growth from 6.8% to 1.1%, the president of India called the numbers "baffling" and the then governor of the Reserve Bank of India (RBI), D Subbarao, called them "analytically bewildering", especially with revisions being made for the same month's IIP over time.[2] Finally, questions have been raised about the validity of the methods used to compile data. Growth numbers released monthly by the Central Statistical Office (CSO) and the annual series released by the *Economic Survey of India* are not in sync. This is because the quality of primary production data supplied by the department of industrial policy and promotion has reportedly deteriorated over time despite a base-year revision (from 1999 to 2004) and expanded coverage of the companies and sectors surveyed (Nagaraj 1999a).

Our aim in this article is to fix the first two problems by reducing the time lag and volatility and to help with the third by reducing the effect of random measurement errors. We can overcome the problem of the IIP being for the past by forecasting it for coming months. Forecasting macroeconomic variables like the IIP is a challenging yet popular exercise in understanding economic growth, given its policy relevance. Industrial production has been widely forecasted using both univariate and multivariate methods, and although the latter are generally considered more powerful, they may "overfit" when the underlying data is noisy.[3]

We seek to provide a forward-looking economic indicator for monitoring industrial sector growth based on the IIP data, which is also robust against the noise in it. We do so in two steps. First, we use simple transformations to ensure the series becomes stationary (that is, not dependent on time). Second, we use univariate modelling on the transformed data. Indeed, an auto-regressive model of order one suffices in our tests to obtain robust estimates six months out, that is, for the two past months for which IIP data are not yet available and for four months into the future. The premise that supports the objective in this article is that even though IIP

numbers themselves may be too volatile on account of measurement and sampling errors, forecasts based on them may be more robust against these errors.[4]

The policy context to this is that at present the share of manufacturing in India's GDP has remained stagnant at 15% to 16% since the 1980s, well below the levels of 25% to 40% in other Asian economies such as China, Malaysia and Thailand. In view of this, the Government of India formulated the national manufacturing policy in 2011 to enhance the share of manufacturing in total GDP to 25% by 2025 and also create 100 million jobs. However, the trend of the proportion of manufacturing to national output has since been in decline – it fell from 15.7% in the fiscal year 2011-12 to 15.2% in 2012-13, and was expected to drop below 15% in 2013-14.[5] With a national manufacturing policy aimed at expanding industrial output, it is critical to monitor the output of the manufacturing sector. The IIP is thus the most important index with the highest weight assigned to its manufacturing sub-series (more than 75%).

Section 2 of the article describes our methodology and Section 3 presents the key results. Section 4 discusses the implications of these results, before we conclude in Section 5.

2 Methodology

The Data

We obtained monthly data series of IIP sub-indices – manufacturing, mining and quarrying, and electricity – as well as the weights to combine them from the official website of the CSO, Ministry of Statistics and Programme Implementation. The present series of the IIP with 2004-05 as the base year is a weighted index indicating industrial output across manufacturing, mining, and electricity with weights of 75.527%, 14.157% and 10.316%, respectively.

For projecting each of the three IIP sub-series, we divided the sample period into an in-sample/initialisation set (January 2006 to April 2012) for modelling and an out-of-sample/test set (May 2012 to May 2013) to subsequently check the quality of forecasts. We tried different subsets of the data and found that a five-year period

provided robust results. Our raw data exhibited seasonality as well as trend, both prominent features of such macroeconomic time series.

Methods Used

The data on industrial production is characterised by seasonal and trend components (Morales et al 1992; Bruno and Lupi 2003a, 2003b; Bulligan et al 2010; Biswas et al 2010; Bordoloi et al 2010; Mazumder and Chakraborty 2013). An issue with this is that seasonal variations could be misinterpreted as trends in the economy. On the other hand, seasonal and trend-adjusted estimates reveal data movements that may otherwise remain hidden.

Therefore, we applied two successive transformations – y-o-y percentage growth to get rid of seasonality, and single-period differences growth to get rid of the trend for each of the three sub-series. The data series thus transformed appears to be random noise and statistical tests confirm stationarity. Next, we used simple univariate models rather than complex multivariate ones to avoid "over-fitting" the data (Makridakis, Wheelwright and Hyndman 1998) and to meet our goal of robust prediction. An overfitted model would not necessarily make good predictions (Frechtling 2001), and might not be robust in the sense of forecasts not being sensitive to the addition of each month's values to the input data series.

After the transformations, we are left with a random component for which univariate forecasting (auto-regressive and moving average or ARMA) is recommended.[6] We first checked for a moving-average (MA) component in the transformed data series and did not find any. Next we tried different levels for autoregression – AR(1) auto-regression with the past month, AR(2) with the past two months, and so on till AR(12), but found that there was significant regression only with the previous month (Bulligan et al 2010; Bagshaw 1987; Bruno and Lupi 2003a; Klose et al 2004; Mayer 2010; Newbold and Granger 1974; Raj et al 2008; Thomakos and Bhattacharya 2005).

We tried other univariate methods as well as multivariate methods. Consistent with Newbold and Granger (1974), we

found that auto-regression gives better results over exponential smoothing and related (Holts-Winter) techniques. We also tried multivariate methods, but did not find them better than the much simpler univariate model. We tried the vector error-correction (VEC) model, in addition to the vector auto-regressive (VAR) method. However, none of the three transformed data series – corresponding to manufacturing, mining, and electricity – indicated any stable long-run relationship among the series. As a result, there was no improvement in forecast accuracy over the much simpler AR(1) models (compared with Biswas et al 2010).[7] Granger tests also indicated no significant causality running between the variables in the multivariate set-up with the three sub-series (after transformation), ruling out short-term relationships between them. We also tried dynamic regressions for the three sub-series of IIP against the manufacturing PMI and Indian stock market-based variables such as CNX-Auto and Sensex, but did not find any significant short-term relationships.[8] Thus we have a strong case to support our use of the univariate auto-regressive method over other methods.

3 Results

To remove seasonality and trends in the series (Box and Jenkins 1970), we first sought to identify these. We used regression with trend, trend-squared, and dummy variables for months and found compelling evidence of these for all the three sub-series (R-squared of 90%+ in all cases; detailed results are available from the authors). Therefore we transformed the raw data (Figures 1a, 2a, p 128 and Figure 3a, p 129) in two steps to remove the seasonality and the trend that was present in the three IIP series for mining, manufacturing, and electricity, and made the data series stationary.

Transformation 1

We converted each of the three IIP sub-index series into ratios by taking the y-o-y percentage difference – the difference between this month's figure and the figure for the same month last year divided by the latter – to remove

Figure 1: Manufacturing Time Series – (a) Raw Data; (b) Year-on-Year Percentage Growth; and (c) Single-Period-Differenced Year-on-Year Growth
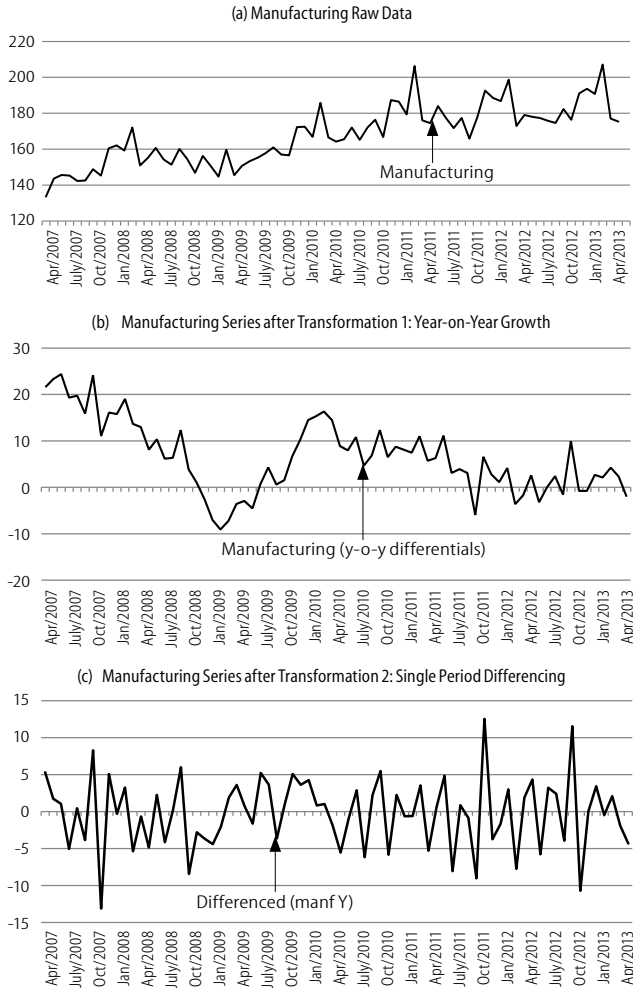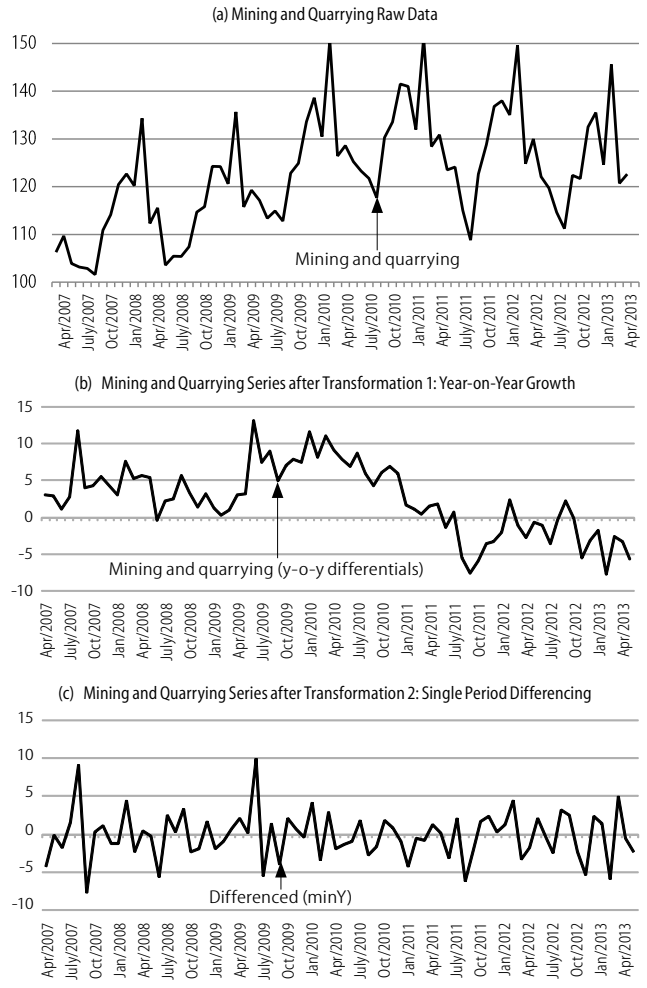

(a) Manufacturing Raw Data

(b) Manufacturing Series after Transformation 1: Year-on-Year Growth

(c) Manufacturing Series after Transformation 2: Single Period Differencing

Figure 2: Mining and Quarrying Time Series – (a) Raw Data; (b) Year-on-Year Growth; and (c) Single Period Differenced Year-on-Year Growth

(a) Mining and Quarrying Raw Data

(b) Mining and Quarrying Series after Transformation 1: Year-on-Year Growth

(c) Mining and Quarrying Series after Transformation 2: Single Period Differencing

seasonality (Figures 1b, 2b and Figure 3b, p 129). On this transformed data, we tested for seasonality using two methods, the Census x12 procedure and regressions of seasonal monthly dummies using the transformed y-o-y growth series. The Census x12 did not show any seasonality for any of the three sub-series using the F-test for stable seasonality; the Kruskal-Wallis Chi-squared test for stable seasonality; or the F-test for moving seasonality. While the regression test showed the continued presence of trend and trend-squared, we found no seasonality after this first transformation.

Transformation 2

To remove any trend, we carried out single-period differencing on y-o-y-transformed series, that is, taking the difference between the figure of this month and the previous month. We tested

the transformed data for stationarity in two ways. First, as before, we used regression and found that neither the trend nor the trend-square components were of any significance. Then we used the Augmented-Dickey Fuller test values to confirm stationarity in the three transformed sub-series (detailed results can be obtained from the authors).

Table 1: Results of Auto-Regressive Model on the Transformed Data

	Manufacturing	Mining and Quarrying	Electricity
Constant	-0.293	-0.154	-0.014
AR (1) (p-value)	-0.383(0.002)	-0.347(0.005)	-0.450(0.0002)
R squared	0.146	0.120	0.207
Adjusted R squared	0.132	0.105	0.194
SE of regression	4.332	2.970	3.124
Probability (F statistic)	0.002	0.005	0.000
Durbin-Watson statistic	2.05	2.06	2.23

Having thus obtained a stationary time series – one with no time dependencies – with much of the "signal" already extracted by the two successive transformations on all three series

(Figures 1c, 2c and Figure 3c, p 129), we could estimate the transformed series to forecast future values, ignoring the "noise" in the data but still extracting any residual signal.

Auto-Regressive Modelling

The AR(1) for the transformed data for all three sub-indices is given in Table 1. Notably, the R-square for each of the sub-indices obtained is low; primarily because these series are essentially noise with the "signal" having been extracted via the two transformations earlier.

The goodness-of-fit statistic is reported typically as a percentage root mean square error (%PRMSE) and root mean square percentage error (RMSPE) to evaluate the projection method using one-month-out

Figure 3: Electricity Time Series – (a) Raw Data; (b) Year-on-Year Growth; and (c) Single Period Differenced Year-on-Year Growth
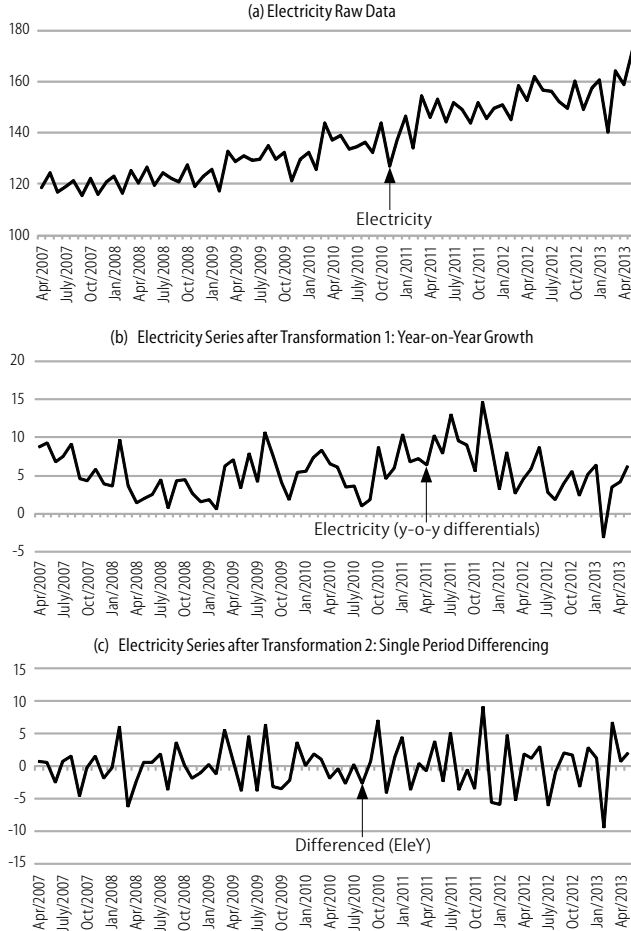
(a) Electricity Raw Data

(b) Electricity Series after Transformation 1: Year-on-Year Growth

(c) Electricity Series after Transformation 2: Single Period Differencing

Figure 4(a): Actual vs Fitted Values for Manufacturing Series
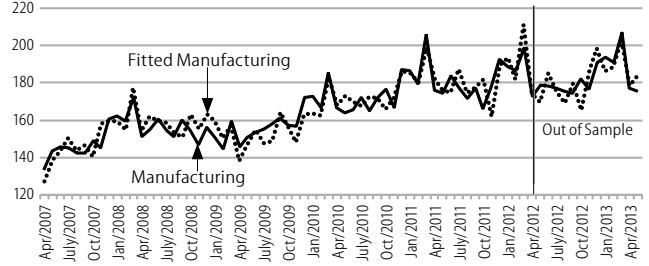(January 2007 to May 2013)

Figure 4 (b): Actual vs Fitted Values for Mining and Quarrying Series
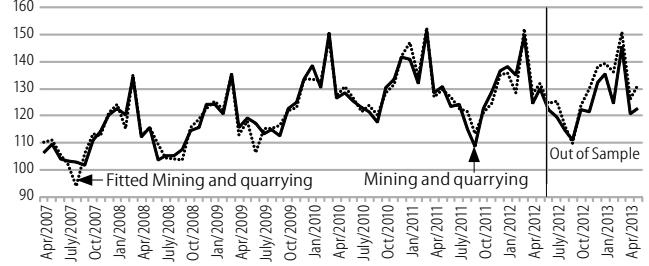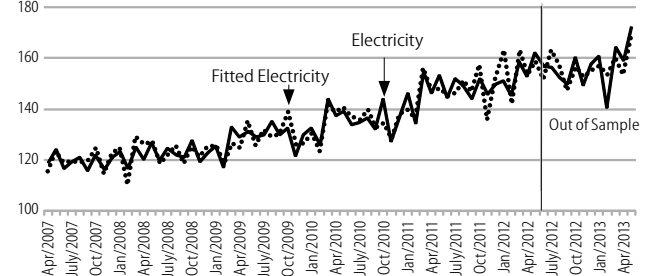(January 2007 to May 2013)

Figure 4 (c): Actual vs Fitted Values for Electricity Series
(January 2007 to May 2013)

forecasts on the out-of-sample/test set (Table 2).

Table 2: Forecast Accuracy Statistics

	Manufacturing	Mining and Quarrying	Electricity	IIP
Percentage root mean square error (PRSME)	4.353	4.473	3.533	3.600
Root mean square percentage error (RMSPE)	4.379	4.637	3.698	3.633

Using this forecasting model and reversing the transformations applied to the time series, we can obtain the actual versus forecasted numbers for manufacturing (Figure 4a), mining (Figure 4b), and electricity (Figure 4c) for both the

in-sample and the out-of-sample periods. A graphical representation of the differenced series is less "spiky", allowing use of the forecast as a robust industrial growth indicator. Moreover, rolling forecasts mean more information for any future month, which can be aggregated into a stable number.

4 Discussion

We want to be able to make projections six months out from the IIP and its sub-series – two months in the past

and four months into the future. We provide six months out projections using AR(1) models on each of the three sub-series of the IIP (Table 3). As a forward-looking index, the six-month projections (or aggregations of forecasts made in subsequent months) provide a basis for policymaking using the future numbers. Indeed, the actual numbers available much later in September 2013 (and still subject to revision) are close to the values we "predicted" in February 2013 when the December 2012 IIP numbers had been just released (Table 3). Note that forecasts for June 2013 can be made in subsequent months as well for a five-month forecast, a four-month forecast, and so on, as we get closer to it, and that these forecasts can be aggregated.

Moreover, as can be expected, our fitted values six months out (forecasts) for all the three sub-series – manufacturing, mining and quarrying, and electricity – have a standard deviation that

Table 3: Six Months Out Forecasts Made with IIP Data up to December 2012 (Available in February 2013) and Subsequent Actual IIP Values (as of September 2013)

	January 2013	February 2013	March 2013	April 2013	May 2013	June 2013
Manufacturing						
Actual value	193.6	190.8	207.1	177	175.4	174.2
Dec 2012 forecast	190.5	188.7	200.7	174.7	180.8	179.9
Mining and Quarrying						
Actual value	135.5	124.6	145.6	120.7	122.6	117.1
Dec 2012 forecast	139.3	136.3	151.1	126.0	131.3	123.3
Electricity						
Actual value	160.7	140.5	164.2	159.1	172.4	157
Dec 2012 forecast	152.7	146.6	160.2	154.3	164.0	158.6

is lower than the actual numbers in the respective periods (Table 4).

Table 4: Standard Deviation (SD) of the Six-Month Projections

Series	Actual Value (SD)	Forecast (SD)
Manufacturing	13.09	9.32
Mining	10.75	10.09
Electricity	10.54	6.17

The smaller standard deviations suggest that noise has been taken out, thereby making the modelling output useful as a benchmark against which the trajectory of growth can be traced. If future IIP numbers appear consistently above our benchmarks for coming months, it could imply improved economic growth. If the actual numbers come out much lower or much higher than our benchmark numbers for coming months, it could imply a one-off number to be safely ignored.

The forecasts are to be made on a rolling basis month after month. The six-month forecast for any given month, say, December 2013 (made in August when the June IIP figures became available), the five-month forecast for that month made in September 2013 using the July numbers, the four-month forecast made in August 2013, and so on should all converge to the actual number when it becomes available. If they do not, a weighted average of previous months' forecasts for this particular month might provide a better estimate than the actual IIP number when it is first announced.

5 Conclusions

We have described how we can project the sub-series corresponding to manufacturing, mining and electricity into the future, and hence overall industrial output as measured by the IIP to obtain a robust and forward-looking indicator of industrial growth. Using two successive transformations we were able to extract much of the "signal" from the sub-series data to obtain growth rate projections that can serve as a useful indicator for determining how well the economy, in particular the industrial sector, is performing compared to the previous year. It is also easy to apply our method to other IIP-related sub-indices such as the capital goods index.

However, there is room for further research to improve the extent to which

the IIP actually reflects Indian industrial production (Nagaraj 1999a, 1999b) by, for instance, incorporating past annual data from the Annual Survey of Industries. Nonetheless, having a rolling six-months robust forecast every time a monthly IIP number is announced can be useful for industrialists to make investment decisions. Policymakers can also find it useful to compare forecasts made against actual numbers to see whether the announced IIP numbers are of value.

NOTES

1 The release of IIP figures is linked to higher price volatility in the stock market.

2 See, for instance, http://articles.economictimes.indiatimes.com/2012-04-14/news/31342106_1_iip-data-sugar-production-trade-data.

3 See Bulligan et al (2010).

4 See "IIP Grows 2.4% but Expect No Rate Cuts", *Economic Times*, 12 July 2012.

5 See "Manufacturing Sector's Contribution to GDP May Fall Below 15% in FY14", *Economic Times*, 12 May 2013.

6 See, for instance, "Seasonal Adjustment and Time Series Issues", United Nations Statistics Division, 17 March 2010.

7 The accuracy metric RMSPE (root mean squared percentage error) for VAR(2) and BAR(2) reported by Biswas et al (2010) are 4.3 and 3.6 respectively against a value of 3.6 from our much simpler univariate method.

8 However, Bordoloi et al (2010) report a smaller figure of 1.14 of RMSPE using a multivariate dynamic factor model.

REFERENCES

Bagshaw, M L (1987): "Univariate and Multivariate ARIMA versus Vector Autoregression Forecasting", Working Paper 8706, Federal Reserve Bank of Cleveland.

Biswas, D, S Singh and A Sinha (2010): "Forecasting Inflation and IIP: Bayesian Vector Autoregressive Model", Occasional Paper, 2317, Reserve Bank of India.

Bordoloi, S, D Biswas, S Singh, S Manna and S Saggar (2010): "Macroeconomic Forecasting using Dynamic Factor Models", Occasional Paper, 2319, Reserve Bank of India.

Box, G and G Jenkins (1970): *Time Series Analysis: Forecasting and Control* (San Francisco: Holden-Day).

Bruno, G and C Lupi (2003a): "Forecasting Euro Area Industrial Production Using (Mostly) Business Surveys Data", Munich Personal RePec Archive Paper (MPRA), 42332, Institute for Studies and Economic Analyses, Rome.

– (2003b): "Forecasting Industrial Production and Early Detection of Turning Points", Discussion Paper, 4/03, Department of Economics, Management, Society and Institutions, University of Molise, Campobasso, Italy, available at http://road.unimol.it/bitstream/2192/66/1/ESDP03004.pdf

Bulligan, G, R Golinelli and G Parigi (2010): "Forecasting Industrial Production: The Role of Information and Methods", Bank for International Settlements, *IFC Bulletin*, 33, pp 227-35, available at http://www.bis.org/ifc/publ/ifcb33v.pdf

Frechtling, D C (2001): "Forecasting Tourism Demand: Methods and Strategies", Reed Educational and Professional Publishing, Oxford, Chapter 6.

Klose, C, M Pircher and S Sharma (2004): "Univariate Time Series Forecasting", Term Paper, Department of Economics, University of Vienna, Austria, available at http://homepage.univie.ac.at/robert.kunst/proguni.pdf

Makridakis, S, S C Wheelwright and R J Hyndman (1998): "Forecasting: Methods and Applications, John Wiley, Singapore, Chapters 4 and 11.

Mayer, M (2010): "Forecasting Austrian Industrial Production with Seasonal ARIMA models in R", University of Vienna, 20 June, available at http://homepage.univie.ac.at/robert.kunst/seas2010_pres_mayer.pdf

Mazumder, S and S Chakraborty (2013): "Study on Seasonal Adjustment of IIP Using X-12 ARIMA", *Journal of Industrial Statistics*, 2 (1), pp 37-67.

Morales, E, A Espasa and M L Rojo (1992): "Univariate Methods for the Analysis of the Industrial Sector in Spain", *Journal of Investigaciones Economicas*, 16 (1), pp 127-49.

Nagaraj, R (1999a): "How Good Are India's Industrial Statistics? An Exploratory Note", *Economic & Political Weekly*, 36 (6), pp 350-55.

– (1999b): "How Good Are India's Industrial Statistics? CSO's Comments", *Economic & Political Weekly*, 34 (23), pp 1461-63.

Newbold, P and J Granger (1974): "Experience with Univariate Time Series and Combination of Forecasts", *Journal of the Royal Statistical Society*, Series A (General), 137 (2), pp 131-65.

Raj, J and S Dhal (2008): "The Inflation Rate in India: Some Applied Issues", Occasional Paper, 29(2), Reserve Bank of India, 47-80.

Thomakos, D D and P S Bhattacharya (2005): "Forecasting Inflation, Industrial Output and Exchange Rates: A Template Study for India", *Indian Economic Review*, 40 (2), pp 145-66.

United Nations Statistics Division (2010): "Seasonal Adjustment and Time Series Issues", 17 March, available at unstats.un.org/unsd/industry/meetings/eclac2011/AC229-20.PPS

Web Exclusives

EPW has introduced a new section, "Web Exclusives" on its new and improved website (http://www.epw.in).

This section will feature articles written exclusively for the web edition and will normally not appear in the print edition. All visitors to the website can read these short articles written mainly on current affairs.

Readers of the print edition are encouraged to visit the EPW website and read these web exclusives which will see new articles every week.