City Research Online

City, University of London Institutional Repository

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

A Robust and Forward-Looking Industrial Production Indicator for India

ManMohan S. Sodhi[1]

Jasmine Sharma[2]

Sukhmeet Singh[2]

Arvinder Walia[2]

[1]**Cass Business School, City University London**

[2]**Indian School of Business, Mohali**

29th September 2013

Abstract: *Against the backdrop of growing criticism of the Index of Industrial Production (IIP) that provides information only about the past and sometimes fluctuated wildly, we seek 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 these are released. Using data on the IIP's three sub-series – mining, manufacturing and electricity – we seek to isolate the 'noise' from the 'signal' in two steps: (1) We isolate seasonal effects and trend components through successive transformation to extract highly informative signals, and (2) we then use an auto-regressive model on the transformed series to forecast for the near-term growth of Indian industry and its sub-sectors six months out. Thus, given the latest available IIP number in any given month (two months in arrears), anyone can use the model presented in this paper to make 'robust' predictions for the two past months and the four months into the future month-after-month.*

I. Introduction

The present study explores the creation of a forward-looking economic indicator to anticipate and understand industrial growth. In the slower growth prevailing in India in 2013 relative to past years, the importance of monitoring industrial growth has only increased. The question is how to do so. The various macro-economic indicators for monitoring manufacturing sector (in particular) are Gross Domestic Product (GDP), Purchasing Managers Index (PMI) and Index of

1

Industrial Production (IIP) however the GDP numbers appear only annually and PMI is essentially an expectations-based index. Therefore, the IIP, which appears every month albeit in arrears by two months, has become the most widely used macro-economic 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%. Policy makers, statisticians, economists, analysts, planners and business entities await and value numbers[1] as indicative of growth in industrial output and make decisions and pronouncements accordingly.

However, doubts have been over the IIP's relevance from a policy viewpoint with at least three problems. *Firstly*, there is a substantial time lag involved i.e. the preliminary numbers are available with a time lag of six weeks from the reference month; thus making these IIP numbers "retrospective". *Secondly*, there are serious problems with the index as regards 'noise' and the 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 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, Dr. D Subbarao, called the IIP numbers "analytically bewildering" especially with revisions being made for the same month's IIP over time.[2] *Finally*, questions have been raised regarding the validity of the data compilation methods. Growth numbers released monthly by the Central Statistics Organization and the annual series released by 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 base-year revision (from 1999 to 2004) and the expanded coverage of companies and sectors surveyed (Nataraj, 1999).

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

As such, we seek to provide a forward-looking economic indicator for monitoring industrial sector growth based on the IIP data that is also robust against the noise in the IIP data. We do so in two steps: (1) We first use simple transformations to ensure the series becomes stationary (i.e., not dependent on time) and (2) we use univariate modeling on the transformed data. Indeed, an auto-regressive model of order one suffices in our tests to obtain robust estimates six months out, i.e., for two past months for which IIP data are not available yet and for futue four months into the future. The premise that supports the objective of our effort in this paper is that even though IIP numbers themselves may be too volatile[4] on account of measurement and sampling errors, forecasts based on them may be more robust against these errors.

The policy context to this is that at present the share of manufacturing in India's GDP has remained stagnant at 15-16% since the 1980s, well below the levels of 25-40% in other Asian economies like China, Malaysia and Thailand. In view of this, the Indian Government of India formulated the *National Manufacturing Policy* in 2011 to enhance the share of manufacturing in total GDP to 25% by 2025 to create 100 million jobs by 2025 (Department of Commerce and Industry, Government of India). However, the trend for the proportion of manufacturing to national output has since been a decline: it fell from 15.7% in 2011- 2012 fiscal year 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 far the most important index with the highest weightage assigned to its manufacturing sub-series (over 75%).

The paper is organized as follows: Section II describes our methodology and Section III presents key results. Section IV discusses the implications of our results before we conclude in Section V.

II. Methodology

The data: We obtained monthly data series of IIP sub-indices, namely manufacturing, mining & quarrying and electricity, as well as the weights to combine them were obtained from the official website of Central Statistics Office (CSO), Ministry of Statistics and Program Implementation (MOSPI). The present series of IIP with 2004-05 as the base year is a weighted index indicating industrial output across manufacturing, mining and electricity with respective 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/initialization set (January 2006 to April 2012) for modeling and an out-of sample/test set (May 2012 to May 2013) to check the quality of forecasts subsequently. We tried different subsets of the data and found that a *five-year period* provided robust results. Also our raw data exhibited seasonality as well as trend, both prominent features of such macro-economic time series.

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

Therefore, we first applied two successive transformations – year-on-year percentage growth to get rid of seasonality and then 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 models to avoid "over-fitting" the data (Makridakis, Wheelwright and Hyndham, 1998) and given our goal for robust prediction. An overfitted model would not necessarily make good predictions (Frechtling, 2001), nor might it be robust in the sense of forecasts not being overly 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 moving-average (MA) component in the transformed data series and did not find any. Next we tried different levels for auto-regression – AR(1) auto-regression with the past month, AR(2) with the past two months, etc. till AR (12) but found that there was significant regression only with the previous month (see also Bulligan et al, 2010; Bagshaw M.L 1987; Bruno & Lupi, 2003; Klose et al, 2004; Mayer 2010; Newbold & Granger, 1974; Raj et al. 2008; Thomakos &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 or electricity respectively – 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] The Granger Tests we conducted 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 Purchasing Managers Index (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 univariate auto-regressive method in comparison to other methods.

III. 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 available from authors). Therefore we first transformed the raw data (Figure 1a, 2a, and 3a) to remove (1) the seasonality and (2) the trend that was present in the three IIP series for mining, manufacturing and electricity respectively to make the data series stationary in two steps:

Transformation 1: We converted each of the three IIP sub-index series into ratios by taking year-on-year (YoY) percentage difference – the difference between this month's figure and the figure for the same month last year divided by the latter – to remove seasonality (Figure 1b, 2b and 3b). On this transformed data, we tested for seasonality using two methods, 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 F-test for stable seasonality; Kruskal-Wallis Chi-squared test for stable seasonality; or F-test for moving seasonality. And, while the regression test showed the continued presence of trend and trend-squared, we found no seasonality after this first transformation.

Transformation 2: Next, to remove any trend, we carried out single-period differencing on YoY-transformed series, i.e., taking the difference between the figure of this month and the

previous. We tested the transformed data for stationarity in two ways. First, as before, we used regression and found that neither the trend not 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 obtainable from the authors).

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 3c), we can use now estimate the transformed series to forecast future values ignoring the 'noise' in the data but still extract any residual signal.

Figure 1. Manufacturing Time Series (a) Raw Data (b) Year-on-Year percentage growth (c) single-period-differenced Year-on-Year Growth

(a) Manufacturing raw data
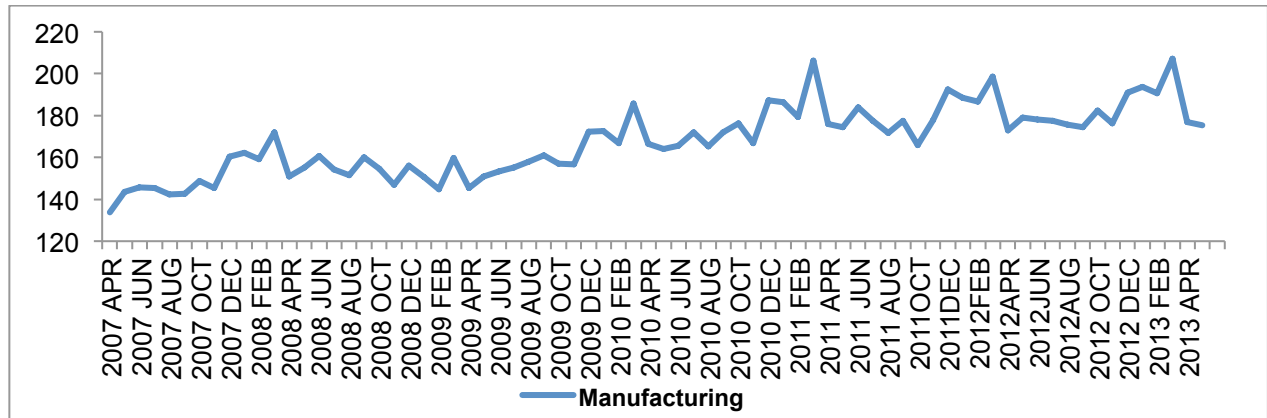
(b) Manufacturing series after Transformation 1: Year-on-Year Growth
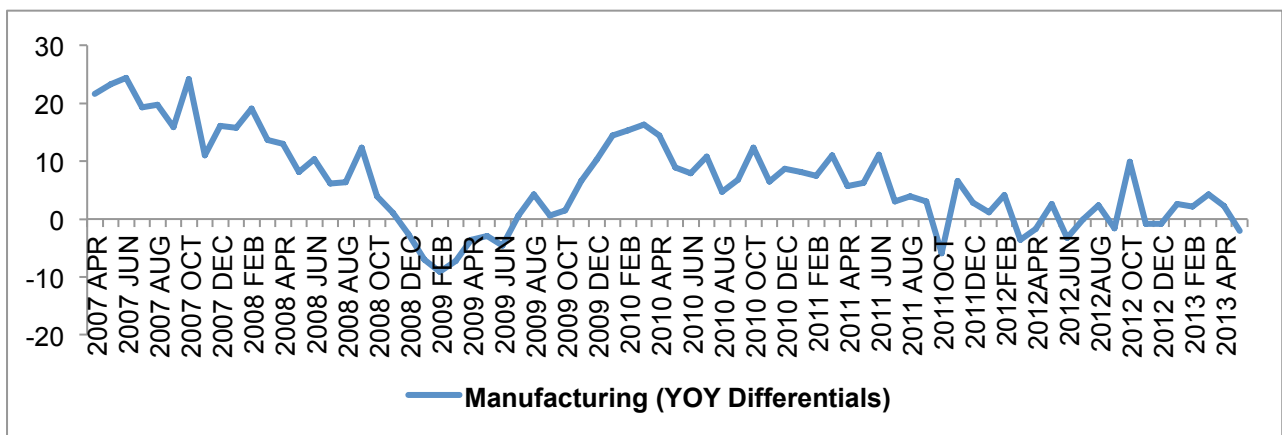
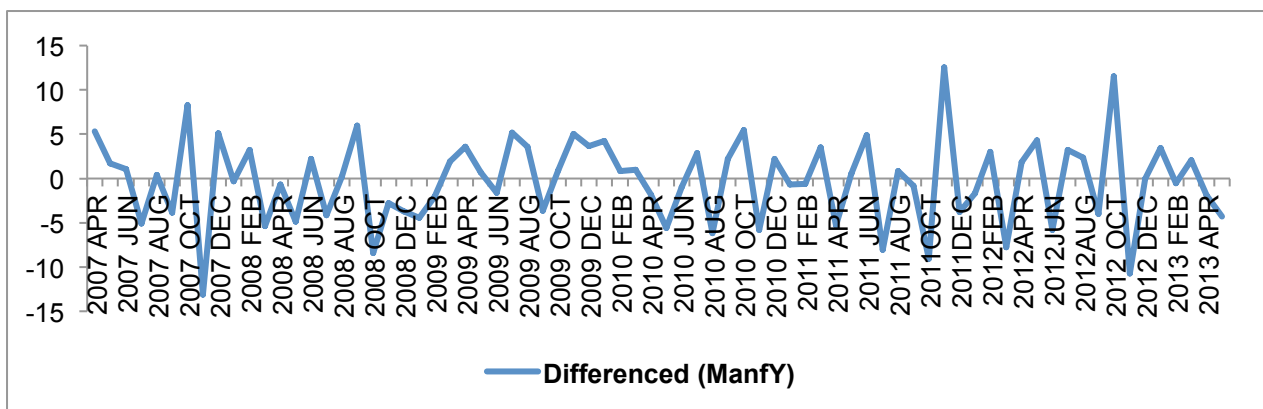**(c) Manufacturing series after Transformation 2: Single Period Differencing**
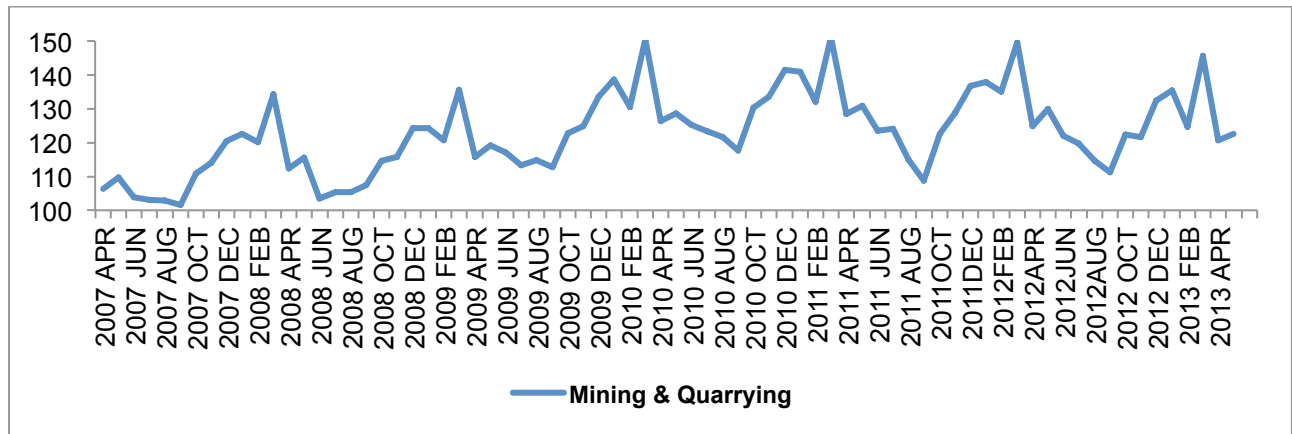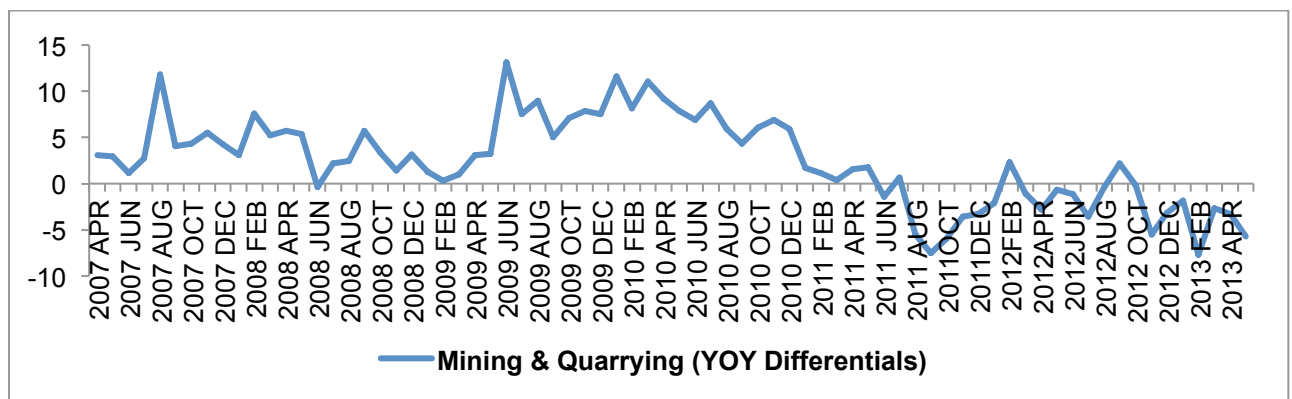

7

Figure2. Mining & Quarrying Time Series (a) Raw Data (b) Year-on-Year Growth (c) Single Period Differenced Year-on-Year Growth

(a) Mining & quarrying raw data

(b) Mining & quarrying series after Transformation 1: Year-on-Year Growth

(c) Mining & quarrying series after Transformation 2: Single Period Differencing
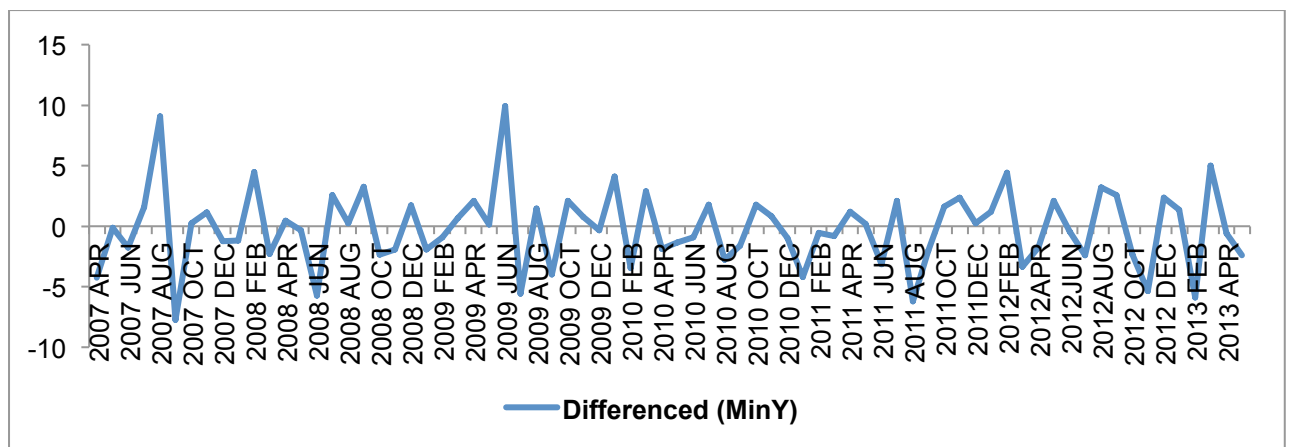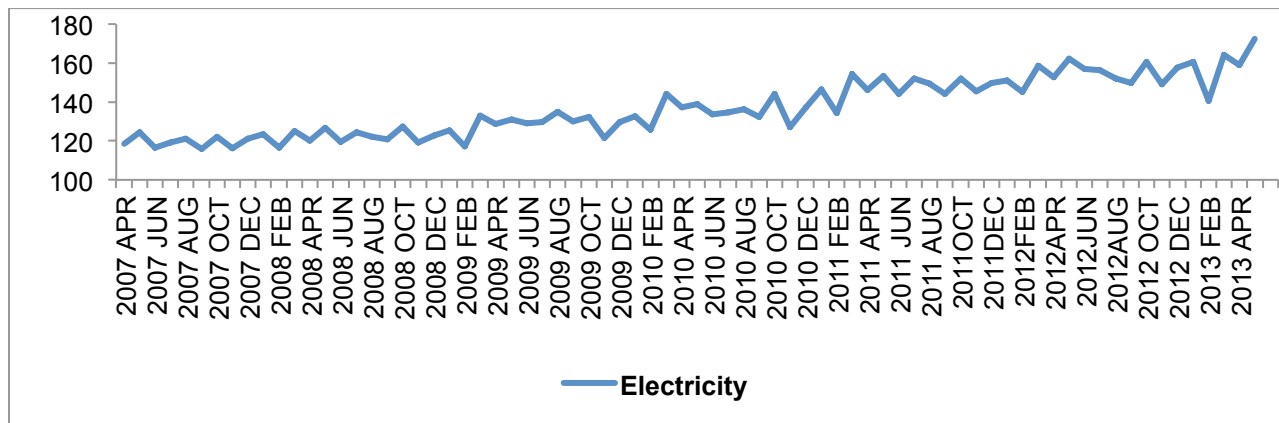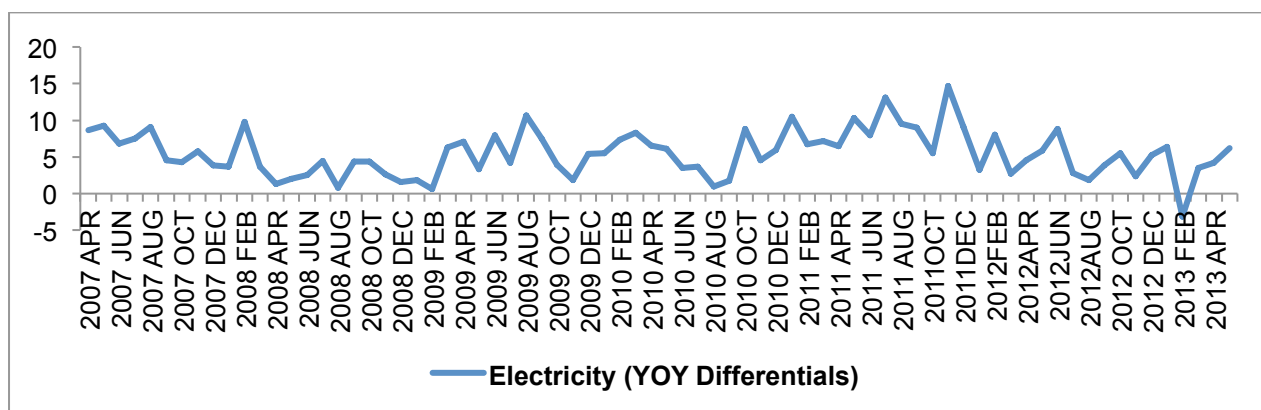
Figure 3. Electricity Time Series (a) Raw Data (b) Year-on-Year Growth (c) Single Period

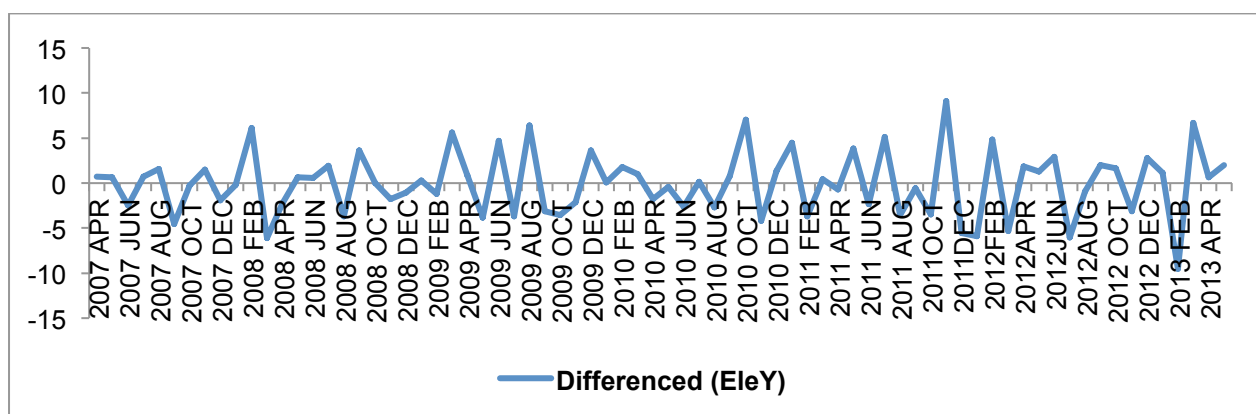**Differenced Year-on-Year Growth**

(a) Electricity raw data

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

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

Auto-Regressive Modeling: 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.

Table 1: Results of Auto-Regressive Model on the transformed data

|	MANUFACTURING	MINING & 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
S.E. 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

The goodness-of-fit statistic is reported typically as percentage root mean square error (%PRMSE) and root mean square percentage error (RMSPE) to evaluate the projection method using one-month-out forecasts on the out-of-sample/test set (Table 2).

Table 2: Forecast Accuracy Statistics

	Manufacturing	Mining & 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' to allow use the forecast as a "robust" industrial growth indicator. Moreover, rolling forecasts mean more information for any future month that can be aggregated into a stable number.
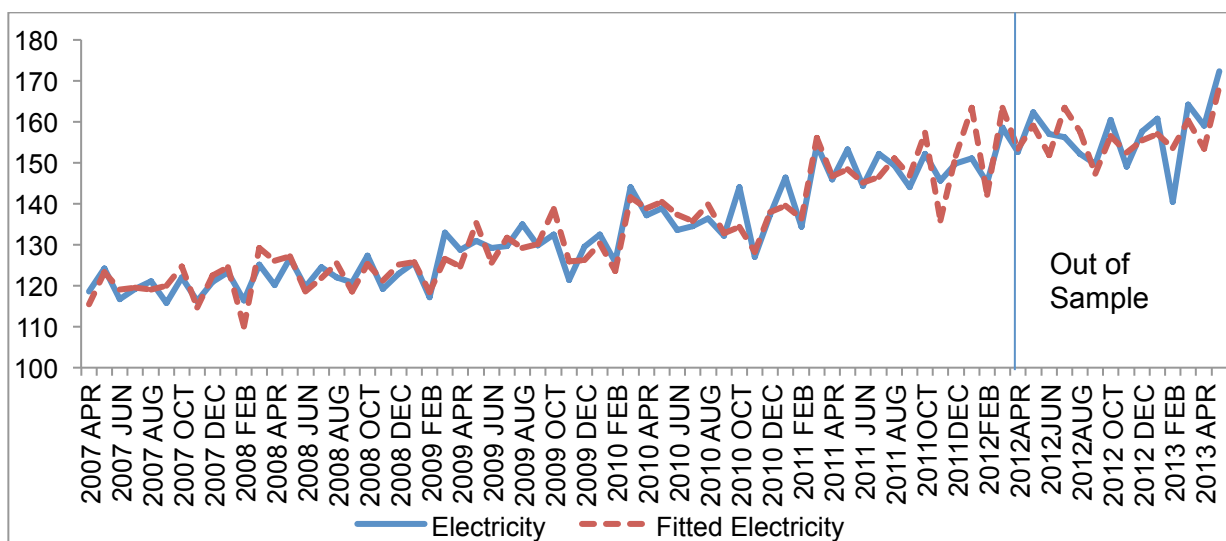
Figure 4: (a) Actual vs. fitted values for manufacturing series (January 2007 to May 2013)
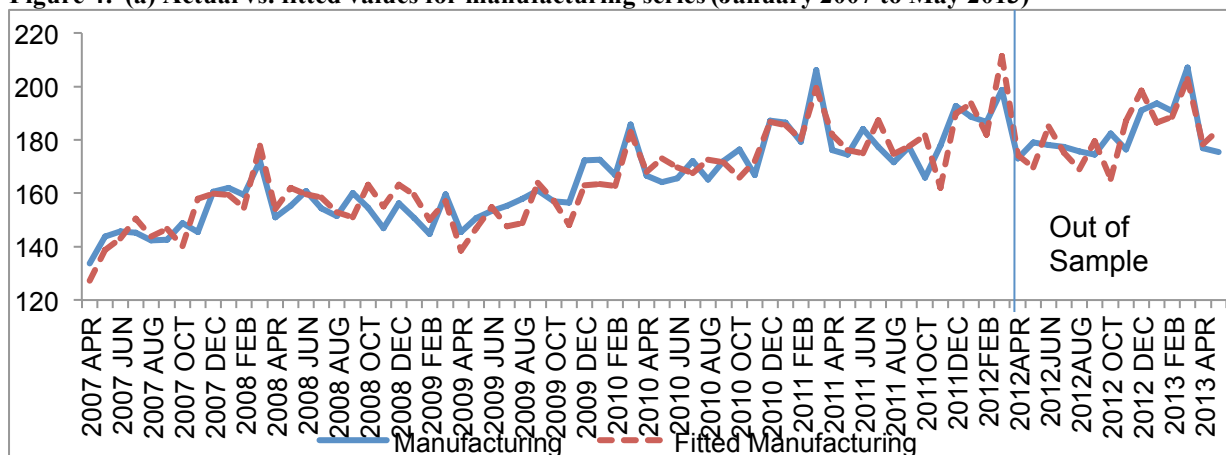
(b) Actual vs. fitted values for mining & quarrying series (January 2007 to May 2013)
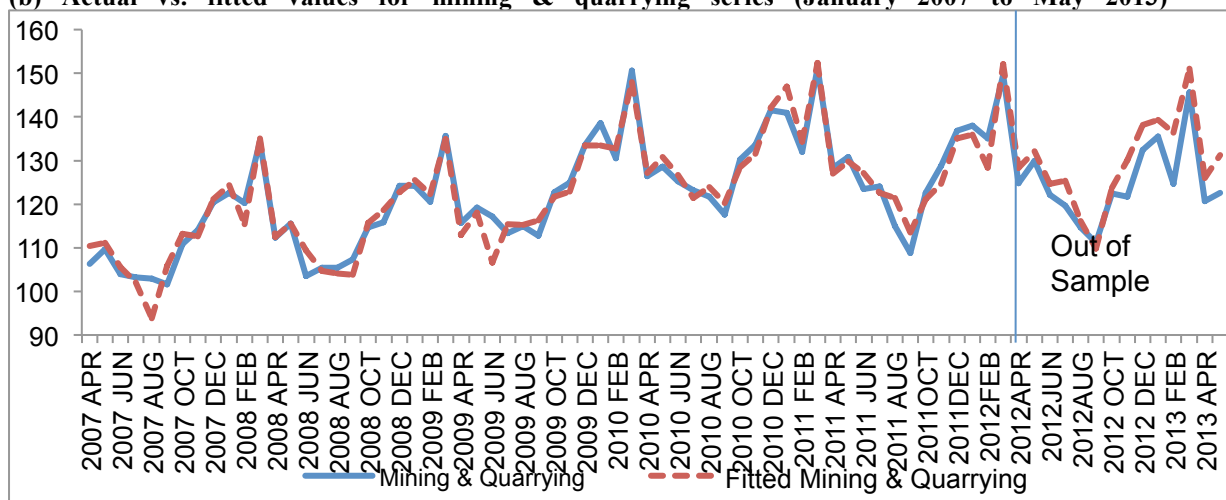
(c) Actual vs. fitted values for electricity series (January 2007 to May 2013)

IV. 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 IIP (Table 3). As a forward looking index, the six-months projections (or aggregations of forecasts made in subsequent months) provide a basis for policy making 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'13 can be made in subsequent months as well for a five-month forecast, a four-month forecast, etc. as we get closer to June'13 with each passing month and these forecasts can be aggregated.

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)

	Jan-13	Feb-13	Mar-13	Apr-13	May-13	Jun-13
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 & 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*

Moreover, as can be expected, our fitted values six-months out (forecasts) for all the three sub-series namely manufacturing, mining &quarrying, and electricity have standard deviation that 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 modeling output useful as a "benchmark" against which the trajectory of growth can be traced. If the 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.

Moreover, 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 June IIP figures become available), the five-month forecast for that month made in Sep'13 additionally using July numbers, the four-month forecast made in August'13, etc., should all 'converge' to the actual number when it becomes available – if they don't, 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.

V. 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 is performing as compared to the previous year as regards to the industrial sector. It is also straightforward to apply our method for 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 (or not) Indian industrial production – see the arguments by Nagaraj (1999) – 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 in their investment decisions. Policy makers can also find it useful to compare forecasts made previously against actual numbers to see whether the announced IIP numbers are of value.

REFERENCES

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

Basu, B (1987): "Measuring Industrial Growth: A Researcher's Point of View", *Economic and Political Weekly,* 22(11), 446-447, March 14.

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.

Bhattacharya, R, R Pandey and G Veronese (2011): "*Tracking India Growth in Real Time",* Working Paper 90, National Institute of Public Finance and Policy, New Delhi.

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, Italy.

_____ (2003b): "*Forecasting Industrial Production and Early Detection of Turning Points",* Discussion Paper, 4/03, Department of Economics, Management, Society and Institutions, University of Molise, 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 (Chapter 57), ISI Session, August 2009, 33, 227-235, available at http://www.bis.org/ifc/publ/ifcb33v.pdf.

Economic and Political Weekly (2011). "The New IIP", *Editorial, Economic and Political Weekly*, 46 (26-27), June 25.

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

Makridakis, S, S C Wheelwright and R J Hyndman (1998): "*Forecasting: Methods and Applications,* (3rd ed.)", John Wiley & Son Ltd, Singapore, (Chapter 4 & 11).

Mayer, M (2010): "*Forecasting Austrian Industrial Production with Seasonal ARIMA models in R.*", University of Vienna, Austria, June 20, 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", *The Journal of Industrial Statistics,* 2(1), 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), 127-149.

Nagaraj R (1999a): "How good are India's Industrial Statistics? An Exploratory Note", *Economic and Political Weekly,* 36(6), 350-355, February6-12.

_____ (1999b): "How good are India's Industrial Statistics? CSO's Comments" *Economic and Political Weekly,* 34(23), 1461-1463, June 5-11.

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), 131-165.

Revilla, P (2002): "*Spanish Methods to Improve Timelessness of in the Industrial Production Indices,* Organization for Economic Cooperation and Development Publication, available at http://www.oecd.org/std/30042667.pdf

Thomakos, D D and P S Bhattacharya (2005): "Forecasting Inflation, Industrial Output and

Exchange Rates: A Template Study for India", *The Indian Economic Review,* 40(2), 145-166.
United Nations Statistics Division (2010): *"Seasonal Adjustment and Time Series Issues", 17 March 2010,*
    available at unstats.un.org/unsd/industry/meetings/eclac2011/AC229-20.PPS.

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-
[3] Bulligan G, Golinelli R , Parigi G; 2010; Forecasting Industrial Production: the role of Information and Methods

[4] Economic Times, 12th July, 2012. IIP grows 2.4% but expect no rate cuts.

http://articles.economictimes.indiatimes.com/2012-07-13/news/32649901_1_industrial-output-industrial-production-

economist-at-care-ratings accessed 18th Oct, 2012.

[5] Economic Times, 12th May 2013. Manufacturing sector's contribution to GDP may fall below 15% in FY14. Available at http://articles.economictimes.indiatimes.com/2013-05-12/news/39203854_1_assocham-survey-100-million-jobs-manufacturing-sector accessed 29th Sep 2013.

[6] See, for instance, United Nations Statistics Division 17th March 2011, " Seasonal Adjustment in Time Series Issues"

[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.