Nowcasting GDP Growth Using a Coincident Economic Indicator for India

Bhadury, Soumya and Ghosh, Saurabh and Kumar, Pankaj

Reserve Bank of India

12 September 2019

Nowcasting GDP Growth Using a Coincident Economic Indicator for India[1]

Soumya Bhadury[1] *Saurabh Ghosh[2]* *Pankaj Kumar[3]*

Abstract

In India, the first official estimate of quarterly GDP is released approximately 7-8 weeks after the end of the reference quarter. To provide an early estimate of the current quarter GDP growth, we construct a Coincident Economic Indicator for India (CEII) using 6, 9 and 12 high-frequency indicators. These indicators represent various sectors, display high contemporaneous correlation with GDP, and track GDP turning points well. While CEII-6 includes domestic economic activity indicators, CEII-9 combines indicators on trade and services along with the indicators used in CEII-6. Finally, CEII-12 adds financial indicators to the indicators used in CEII-9. In addition to the conventional economic activity indicators, we include a financial block in CEII-12 to reflect the growing influence of the financial sector on economic activity. CEII is estimated using a dynamic factor model to extract a common trend underlying the high-frequency indicators. We use the underlying trend to gauge the state of the economy and to identify sectors contributing to economic fluctuations. Further, CEIIs are used to nowcast GDP growth, which closely tracks the actual GDP growth, both in-sample and out-of-sample.

JEL Classification: C32, C51, C53

Keywords: *Nowcast, Gross Domestic Product, Economic Cycle, Dynamic Factor Model, Turning Point Analysis, Jagged Edge Data*

[1] Soumya Bhadury is Manager-Research in the Strategic Research Unit (SRU), RBI. [2]Saurabh Ghosh is Director in SRU, RBI. [3] Pankaj Kumar is Director in SRU, RBI. The views and opinions expressed in this paper are those of the authors and do not necessarily represent those of RBI. The authors thank Prof. Pami Dua, Dr. Jay Surti and an anonymous referee for their valuable comments on an earlier draft of the paper. Research assistance provided by Twinkle Roy and Ayushi Jain are gratefully acknowledged.

I. Introduction

Real-time assessment of the state of the economy is a pre-requisite for making appropriate policy decisions. The effectiveness of policy-making depends on how well it uses all the information available at any given time. Moreover, making use of the current flow of information is an essential ingredient in formulating a forward-looking policy. Today, in a dynamic environment, data driven policy making is in vogue. This is all the more important because of significant lags in official data releases of key macro variables such as GDP – an all-encompassing measure of economic activity. Emerging market economies face serious problems of data lags, gaps and revisions which hamstring optimal policy decisions.

A core concern in policymaking is identifying the signs of expansions and contractions in economic activity. At any point, diverse economic indicators may indicate varied trends in activity. Therefore, combining all of these together in an appropriate way to arrive at the underlying (or unobserved) trend has traditionally occupied the attention of both governments and businesses. Traditionally, a lot of work was done in the National Bureau of Economic Research and the Department of Commerce in the US on identifying business cycles using multiple indicators. Often, their analysis relied on prior beliefs and judgement. Later, academic interest evolved in this area. It contributed by making the analysis more formal and mathematically precise (see Stock and Watson, 1989). Therefore, presently this whole exercise displays a mix of both sound conceptual framework and careful inference and judgement.

There exists a two-pronged challenge in assessing the underlying state of the economy using high frequency indicators. The first one is the choice of appropriate indicators from a large set of potential indicators. The second challenge is associated with signal extraction from the chosen indicators. This is because the individual indicators may reflect short-term idiosyncrasy rather than an underlying general trend. Researchers look at a variety of indicators relating to different aspects of the economy – production, income, sales and employment – together to assess the underlying state of the economy. The consensus trend emanating from all these indicators suggests recessions or expansions (or equivalently slowdowns and accelerations) in economic activity. The coincident economic indicator tries to address the challenge of signal extraction by identifying the general (or common) trend underlying several activity indicators.

To address the first challenge of identifying appropriate indicators, our broad approach is to weigh indicators based on their information content about the dynamics of GDP. We identify a pool of relevant indicators based on 1) contemporaneous correlations between indicators and GDP and 2) explanatory power of the indicators around GDP turning points. Based on these two criteria, we choose a set of twelve indicators which includes index of industrial production – consumer goods and core infrastructure, automobile sales, non-oil non-gold imports, exports, rail freight, air cargo, foreign tourist inflows, government tax receipts, Sensex, NEER and bank credit. Despite following an independent variable selection approach, our choice of indicators is similar but not the same as Stock and Watson (1989) which used index of industrial production, real personal income, real manufacturing and trade sales, and employee-hours in non-agricultural establishments to construct a Coincident Economic Indicator. One reason for deviating from Stock and Watson (1989) is that in emerging markets such as India, indicators on employment and income are hard to obtain at high-frequency. Therefore, researchers have used proxy variables as the second-best option. The second reason for deviating from the indicator list of Stock and Watson (1989) is due to unavailability of such indicators over a long sample. On the whole, our choice of indicators represent real, external and financial sectors which are used to extract the common underlying trend using a single-index dynamic factor model. We call the estimated index the coincident economic indicator for India (CEII).

Given the lags in the release of official GDP, it is popular in the literature to use the estimated coincident economic indicator to nowcast GDP. Broadly, the nowcast exercise is an extension of the bridge type regressions that relate GDP to quarterly aggregates of a few relevant economic indicators. However, we rely on a parsimonious autoregressive model of GDP augmented by CEII to nowcast current quarter GDP. While tracking the actual GDP dynamics closely in the sample, our modelling approach also shows considerable gains in terms of out-of-sample performance.

The rest of the paper is organized as follows. Section II provides a comprehensive review of the literature on coincident indicators and nowcasting, including cross-country studies as well as literature specific to India. Section III briefly describes the data used for constructing CEII. Section IV presents methodology and empirical results on indicator selection, estimation of a dynamic factor model, nowcasting GDP and performance evaluation of our models. We put our concluding remarks in section V. Appendix contains technical details.

II. Literature

Nowcasting involves an exercise of predicting the present, the very near future, and the very recent past- and that makes it more effective in shorter horizon forecasting (Banbura, Gianonne and Reichlin, 2010). A set of coincident indicators is commonly used in the GDP nowcasting exercise. Perhaps, the first use of coincident indicators is found in Burns and Mitchell (1946), which popularized the study of business cycles, and that eventually led to the creation of the composite index of coincident indicators. The methodology has gradually been refined over time and a breakthrough came in the Stock and Watson's (1989) seminal work in estimating a single-index dynamic factor model (DFM). More recently, Giannone, Reichlin and Small (2008) have pioneered the usage of DFM based on a large number of high-frequency indicators and found that the nowcasts have outperformed standard univariate models like random-walk and autoregressive (AR) models.

The coincident indicators are widely followed in gauging the health of an economy. The indicators used are in-sync with the current economic cycle and primarily represent six blocks which include a) industry and construction, b) personal income (or consumption), c) payroll employment d) services e) external sector and f) price. These indicators are chosen not only because of their mere correlation with GDP, but also because they contain timely information about the target variable.

Depending on countries and sample periods, various researchers have also included miscellaneous economic activity indicators (e.g. air cargo, rail freight, port traffic and uncertainty / volatility indices). Though, most of the financial sector variables (e.g. money supply, currency with public, credit, equity indices and bond indices) are generally classified as leading indicators, some of the studies have incorporated them in GDP nowcasting and documented improvement in their nowcasting performances (Annex Table I and II report the details). Some of the researchers have included financial conditions indices (a linear combination of risk spread, asset returns and volatility) to model the risk to output growth one year into the future (IMF, Global financial Stability Report, October 2017)[2]. More complicated models with rich cluster of variables have been commonly used for GDP forecasting rather than nowcasting, mainly considering lags in data releases and their marginal contribution in GDP nowcasting.

[2] The IMF Financial Stability Report states that emerging economies have enhanced their resilience and improved their macroeconomic outlook of output growth. Lower corporate leverage have prospects of positive output growth spillovers but financial stability could be a concern due to political pressure.

The core nowcasting methodology generally involves three steps. First step includes estimating the underlying factor from a set of high frequency economic indicators in the training sample. Generally, principal component analysis, DFM or Bayesian DFM is used in this stage. Second step involves projecting the common factor in the test sample and the methodology used is Kalman-Filter. Finally, the third step involves using the factor projections in an appropriate autoregressive or bridge equation set-up to nowcast GDP. Researchers have also used *mixed data sampling* or expectation maximization algorithm at this stage, especially for EMEs.

II.1 Cross-country nowcasting models

Modern macro-literature emphasizes the role of forward looking assessment of growth and inflation by central banks in policy formulation. Globally, central banks and their monetary policy committees (MPCs) rely on high-frequency economic indicators for an assessment of the current state of the economy. For instance, Bank of England's MPC uses such model-based nowcast to inform its monthly monetary policy decisions. Similarly, the Federal Reserve Bank of Atlanta (FRBA) [3]and the Federal Reserve Bank of New York (FRBNY)[4] publish their own model-based nowcast in addition to the routine forecast produced in the bank. The European Central Bank[5] also consults a DFM-based nowcasting model to inform its monetary policy decisions. Other central banks such as Norges Bank [6]uses nowcasting models of GDP to inform its policy rate decisions.

Outside the ambit of central banks, several country-specific GDP nowcasts have evolved over time. Among the advanced economics (AEs), Kumar (2013) constructed a high-frequency real activity indicator that tracked economic activity in Canada reasonably well. The indicator is obtained as an unobserved common factor capturing the co-movements in real macroeconomic variables. Similarly, Chikamatsu et al. (2018) produced nowcasts of quarterly GDP estimates for Japan by adopting a bridge equation approach. The bridge equation links the low-frequency variables and the index obtained from the high-frequency indicators. Annex Table I provides an exhaustive list of variables used by the authors.

---

[3] GDPNow forecasting model by FRBA does a "nowcast" of the official estimate prior to its release by estimating GDP growth using a methodology similar to the one used by the U.S. Bureau of Economic Analysis.

[4] The FRBNY nowcast model produces forecasts of GDP growth for all variables taking into account their dynamic interactions.

[5] The European Central Bank model the monthly data as a parametric dynamic factor model cast in a state-space representation against the quarterly GDP.

[6] Norges Bank analyzes unstructured textual information of a business newspaper to decompose daily news topics and nowcast quarterly GDP growth for policy rate decisions.

While DFM based nowcasting models have been adopted for a long time across advanced economies, the emerging market economies (EMEs) have started exploring the usefulness of such models only recently. For instance, Luciani et al. (2015) used high frequency indicators (e.g. cement, PMI-manufacturing, consumer confidence, auto sales, etc.) for Indonesia in estimating a dynamic factor and then used the same to nowcast GDP growth in an autoregressive, bridge-equation framework. Caruso (2015) used the Maximum Likelihood estimation in an expectation-maximization (EM) algorithm for constructing a coincident index for Mexico. The EM algorithm is useful in analyzing non-synchronous data releases that are often observed in EMEs.

II.2 Nowcasting Indian GDP

Among the emerging markets, India has a reasonably long history of research in tracking and nowcasting GDP. A few published works in this domain include Dua and Banerji (2001), and Technical Advisory Group constituted by the Reserve Bank (2002, 2006). Dua and Banerji (2000) published an index of monthly coincident indicators to help ascertain the timing of recession and expansion of economic activities based on a set of objective indicators that are synchronous with cyclical fluctuations in growth. Some of the indicators (e.g. monthly unemployment numbers published in the Monthly Abstract of Statistics) are not published anymore. However, we report an update of this monthly index, using available data and proxy variables, in the Annex Note I.

RBI (2006), on the other hand, published a set of indicators including Composite Index of Leading Indicators (CILI) and Composite Index of Coincidental Indicators (CICI) based on a detailed empirical exercise using growth cycle and growth rate cycle methodologies. CILI and CICI were mainly based on principal component analysis with sample data spanning from April 1990 to March 2006. However, the Group in its recommendations mentioned that there remains some methodological issues which need further strengthening in subsequent policy research. Our construction of a coincident economic indicator for India (CEII) is an attempt to address some of the methodological issues highlighted in the report.

Recent academic research, such as Dalhaus et al. (2017) and Bragoli and Fosten (2017) adopted a standard dynamic factor model (DFM) framework for GDP growth nowcasting. Both these studies attempted to extract monthly unobserved common factor from a set of monthly indicators. Following factor extraction, monthly projection of GDP growth is estimated based on the dynamics of the common factor. The indicators used in these studies are summarized in Annex Table II. Dua and Sharma (2016), on the other hand, used a univariate Markov regime switching model to characterize growth cycle phenomena and distinct economic regimes for India and compared them with the US, UK, Germany and Japan. In addition to the academia and central banks, market research organizations, think-tanks, and professional economists have estimated and published coincident indicators and used them to nowcast Indian GDP. For instance, Rabobank[7] uses different models to nowcast India's GDP growth that include Bayesian VAR (BVAR), OLS and a combined model, which help in deriving the underlying contribution of high-frequency indicators for GDP growth forecasts. Besides economic indicators, Rabobank also includes financial indicators such as monetary base (M_0), volatility index (VIX), BSE-500 index and Sensex.

National Institute of Public Finance and Policy (NIPFP) estimated common factors using PCA and then used three different variants of bridge equation models. The NIPFP working paper used variables from several sectors of the economy that included industrial production, construction, services and financial sectors and documented that the empirical strategy outperformed the benchmark AR models. An NCAER working paper presented a new framework to nowcast India's GVA using information of mixed data frequencies and adding evening-hour luminosity information to capture the economic activities of informal sectors in India. However, these nowcasting exercises have not been subsequently updated.

In India, professional economists (economists with banks, brokerage houses and think tanks) also regularly nowcast GDP using proprietary models. Median of their forecasts is published on the RBI website. Some of the GDP nowcasts by these economists are also published on the Bloomberg[8], however their exact nowcasting methodologies are not in the public domain. Bloomberg also publishes its *Monthly GDP tracker* in providing an advance estimate of the current quarter economic activity in India. The *Monthly GDP tracker* is constructed by applying weights to the monthly activity indicators, such as agricultural trade balance, real currency demand, industrial production, etc. (details in annex) and the weights assigned to each monthly indicator is the inverse of the standard deviation of the respective indicator.

To summarize, the use of coincident indicators in GDP nowcast has a long history in advanced economies and still being actively used in monetary policy making by the central banks. Though there has been debates relating to inclusion of high frequency variables, their lead, lag or coincident characteristics, nowcast exercise has moved on by including new blocks of variables

[7] Rabobank combines predictions from Bayesian Vector Autoregressive (BVAR) model and an Ordinary Least Square (OLS) model for GDP forecasts. Link: https://economics.rabobank.com/publications/2019/february/nowcasting-the-indian-economy/
[8] Bloomberg tracks real time GDP data using a weighted methodology to nowcast GDP from high frequency volume based economic indicators.

with an objective to improve nowcast performance. Application of coincident indicators and GDP nowcasting in economic policy making is relatively new in emerging economies and have been challenged by small sample size, non-synchronous data releases and varying data lags. However, considering its importance in policy, economists have been striving with new empirical strategies to bridge this gap and use coincident indicators as an active policy tool.

III. Data

We intend to construct an index that tracks the business cycle reasonably well. The first best is always the quarterly estimates of GDP at constant prices which is published by the Central Statistical Office (CSO[9]) with a lag of two months. Given the delay in data release and the general criticism that it may not represent a pervasive and pronounced downswing in a variety of measures, we use Coincident Economic Indicator for India (CEII) in providing the current-quarter nowcast of GDP, our target variable. The indicators used in the construction of CEII represent all relevant sectors of the Indian economy. Data published on a monthly frequency is exploited to construct the CEII and they are sourced from the CEIC database. CEIC is a data aggregator that collects data from different ministries, government documents and other data originators.

The high-frequency monthly series associated with industry and construction block is IIP-core. Personal income and consumption block is represented by indicators such as IIP-consumer goods and auto sales. Government tax revenue, exports, and non-oil and non-gold imports represent the services and external sector blocks respectively. The miscellaneous economic activity is represented by rail freight, air cargo and foreign tourist arrivals. In addition to the CEII-6 and CEII-9, we have developed a CEII-12 that adds a credit and finance block by including non-food credit, NEER and Sensex. Though in the literature, equity index and credit in some instances have been classified as lead indicators, we explain the rationale for inclusion of these variables in the CEII-12 in the variable selection section.

At this juncture, it might be important to mention that the GDP series has been revised and rebased in India from time to time. Recently, on November 28, 2018, the Ministry of Statistics and Programme Implementation released the annual back series of data beginning 2004-05 (at 2011-12 prices). As per the new series, GDP growth has been revised downward during 2004-05 to 2011-12 compared to the earlier 2004-05 series. However, the Ministry has not separately released

[9] The Ministry of Statistics and Programme Implementation has decided to merge the CSO and National Sample Survey Office (NSSO) into National Statistical Office (NSO).

the back series of the quarterly data. Therefore, we simply splice the 2011-12 quarterly data backwards using the older 2004-05 and 1999-2000 series. However, as a robustness exercise, we also use the downwardly revised Y-o-Y growth of our quarterly dataset in line with the decline in annual growth suggested by the new back series. This, however, does not change our estimates significantly.

IV. Methodology and Empirical Results

The computation of coincident indicator essentially boils down to variable selection, standardization, smoothing and appropriately combining these transformed variables into an index. This index then is appropriately used for GDP nowcasting. We approach the above in four sequential stages which are described in detail in the following sections.

IV.1 Variable Selection

Variable selection is perhaps the most crucial part of the exercise. Essentially the aim is to include variables that would capture the pronounced and persistent movements in economic activity. We look at the availability of high frequency data over sufficiently long sample and existing literature that are particularly relevant for India e.g. Dua and Banerji (2000), RBI (2002) and RBI (2006). However, the Indian economy has evolved considerably over the last decade. We compute the dynamic correlation coefficients and use a forward step-wise and Lasso selection procedure around the turning points in GDP to select the relevant set of indicators, from a set of 27 high-frequency variables available for the Indian economy.

Dynamic Correlation

We look at the correlation of Q-o-Q seasonally adjusted annualized growth of the variables with our target variable, GDP. The dynamic correlation analysis is carried out over 2003:Q1 to 2019:Q1 with around 64 observations at different leads and lags of GDP. We have identified variables as Coincident ('C'), Leading ('L+') and Lagging ('L-') indicators depending on their contemporaneous, future and past correlations, respectively with GDP. These correlation coefficients and their statistical significance are presented in the Annex Table-III. Among the economic indicators, the highest contemporaneous correlation with GDP is observed for Sensex, automobile sales and air cargo. Indicators, such as non-oil and non-gold imports, personal loans and steel consumption represent leading properties in terms of their forward correlation with GDP. Finally, a few indicators e.g. IIP-consumer goods and IIP-core have displayed both coincident and leading properties in terms of their correlations with contemporaneous and future GDP. The summary table 1(a) below reports the dynamic correlation coefficients for the shortlisted high-frequency indicators.

Table 1 (a): Summary of selected variables using Dynamic Correlation

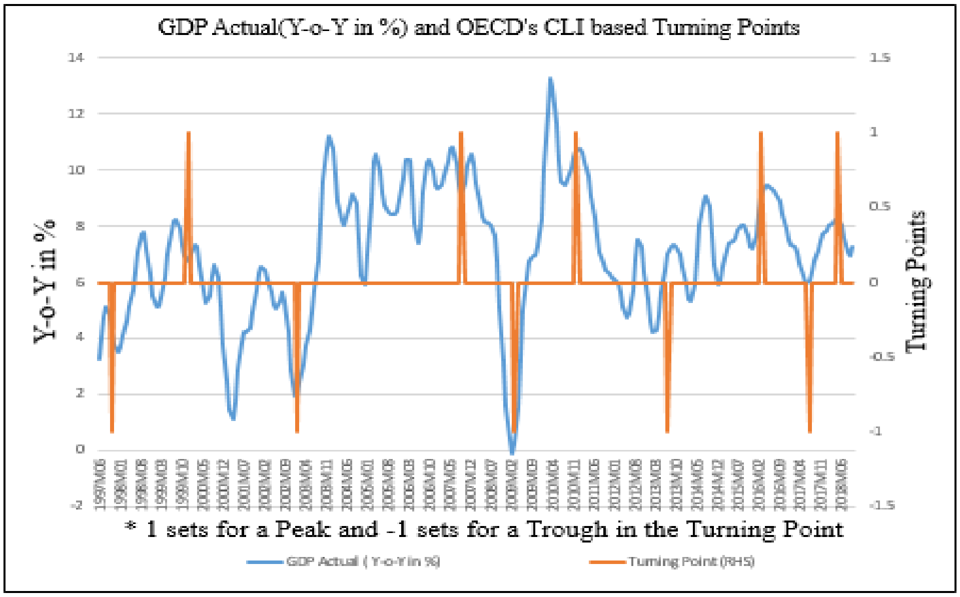
Variables	GDP (t-3)	GDP (t-2)	GDP (t-1)	GDP (t)	GDP (t+1)	GDP (t+2)	GDP (t+3)	Indicator Type
AIR CARGO	0.16	0.01	0.29*	0.46*	0.23	-0.13	-0.12	C
AUTO TOTAL	0.05	-0.09	0.13	0.30*	0.24	-0.06	-0.14	C
BANK CREDIT	-0.09	0.02	-0.06	0.02	0.09	0.23	-0.02	X
EXPORTS	-0.11	-0.31*	-0.08	0.30*	0.41*	0.33*	0.15	C/L+
FOREIGN TOURIST	0.17	-0.07	0	0.34*	0.23	0.04	0.03	C
GOVT. RECEIPT	0.27*	-0.03	0.02	-0.04	0.25*	-0.08	-0.07	L+
IIP CONSUMER	0.07	0.04	-0.15	0.39*	0.52*	-0.07	-0.09	C/L+
IIP CORE	0.26*	-0.26*	0.09	0.30*	0.28*	-0.08	-0.03	C/L+
NONG	0.04	-0.07	-0.11	0.02	0.32*	0.31*	0.30*	L+
NEER	0.09	0.13	0.23	0.26*	0.2	0.12	0.17	C
RAIL FREIGHT	-0.07	-0.12	0.13	0.28*	0.23	-0.31*	-0.15	C
SENSEX	0.02	-0.05	0.23	0.54*	0.41*	0.06	-0.1	C

Note: NONG: Non-oil-non-gold imports; IIP Inf.: IIP infrastructure; IIP Inter: IIP Intermediate

*: Indicates 5 per cent level of significance; C: Coincident, L+: Leading and L-: Lagging indicator

Next, we introduce the forward step-wise selection, followed by the Lasso procedure to identify indicators that are relevant around turning points in GDP. The turning points are identified using the OECD's Composite Leading Indicators (CLI) based "growth cycle" approach. OECD calculates CLI for 33 member countries and 6 non-member countries including Brazil, China, India, Indonesia, Russia and South Africa. OECD uses the turning point detection algorithm, which is a simplified version of the original Bry and Boschan (B&B) routine. B&B routine parses local minima and maxima in the cycle series and applies the censor rules to guarantee alternating peaks and troughs. OECD's CLI based approach has identified 10 such turning points for the Indian economy with +1 set for a peak and -1 set for a trough in the turning point[10]. Figure 1 below plots the real GDP growth and the CLI based turning points around the growth cycle.

Figure 1: Composite Leading Indicator based turning points in the business cycle



Forward Step-wise Selection

The forward stepwise selection procedure is a computationally efficient alternative which sequentially add candidate variables to assess their predictive power around the turning points in GDP. In particular, at each step the variable that gives the highest incremental improvement (to the fit) is included in the model. Based on the forward step-wise selection criterion, a detailed listing of variables that are found to be relevant around turning point in the GDP growth cycles is reported in the Annex Table IV. The summary table 1(b) below provides a relative ranking of the variables in terms of their relevance around turning points. In terms of relative ranking, foreign tourist, NEER and rail freight have highest relevance around the turning points in GDP.

[10] The reference chronology of the turning points are as follows: trough in October-1997, peak in December-1999, trough in January-2003, peak in September-2007, trough in March-2009, peak in December-2010, trough in July-2013, peak in March-2016, trough in July-2017, and peak in May-2018.

Lasso

Lasso shrinks the coefficient estimates towards zero by forcing some of the coefficient estimates to be exactly equal to zero when the tuning parameter, λ is set at a sufficiently large value. Thus, much like the forward step-wise selection procedure, the lasso technique performs variable selection. The lasso coefficient β_λ^L minimize the quantity and $|\beta_j|$ is the lasso penalty

$$\sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij})^2 + \lambda\sum_{j=1}^{p}|\beta_j| \quad = \quad RSS + \lambda\sum_{j=1}^{p}|\beta_j|$$

Lasso yield sparse model, which involves a subset of the variables that makes model interpretation easier. Based on the Lasso criterion, a list of indicators that are found to be relevant around turning point in the GDP growth cycles is reported in the summary table 1(b).

Table 1 (b): Summary of selected variables using Lasso and Forward Step-wise Selection

CEII	Indicator	Dynamic Correlation	Dynamic Correlation	Dynamic Correlation	LASSO	Forward step Wise Selection
		GDP(t-1)	GDP(t)	GDP(t+1)	GDP(t)	Relative Ranking
CEII-6	Air cargo	0.29*	0.46*	0.23	Y	4
	Auto sales	0.13	0.30*	0.24	Y	5
	Rail freight	0.13	0.28*	0.23	Y	3
	Govt. receipt	0.02	-0.04	0.25*	Y	6
	NONG	-0.11	0.02	0.32*	N	11
	IIP consumer	-0.15	0.39*	0.52*	Y	8
CEII-9	Exports	-0.08	0.30*	0.41*	N	9
	IIP core	0.09	0.30*	0.28*	N	12
	Foreign tourist	0	0.34*	0.23	Y	1
CEII-12	Sensex	0.23	0.54*	0.41*	N	10
	Bank credit	-0.06	0.02	0.09	Y	7
	NEER	0.23	0.26*	0.2	Y	2

*: Indicates 5 per cent level of significance; # relative ranking for Forward Step-wise selection procedure; Y: Relevant indicators around turning points in GDP using Lasso, N: Indicators that are not relevant around turning points in GDP.

To summarize, our main objective is to construct a coincident economic indicator for India based on economic activity indicators that co-move strongly with the target variable. Our first round of screening is based on the dynamic correlations between high-frequency indicators and GDP. Further, to refine relevant indicators around turning points, we look at the forward step-wise selection and Lasso criteria. We construct a battery of coincident indicators that include a CEII-6 model covering exclusively domestic economic activity indicators, a CEII-9 model incorporating trade and services sector, and finally a CEII-12 which incorporates financial variables. Furthermore, we equally consider experts' views, past works and judgement in identifying and grouping variables.

IV. 2 Dynamic Factor Model

After selecting the relevant variables to be incorporated in our model, we proceed to estimate a single factor representing the common trend underlying these variables. This is achieved by estimating a dynamic factor model (DFM), a procedure widely popular in the recent literature. For analytical clarity, we sequentially estimate 6-indicators, 9-indicators and 12-indicators DFMs.

Our dynamic factor model contains the following set of equations representing a state-space form.

$$
\begin{aligned}
&x_t = x_{t-1} + w_t \text{ where } w_t \sim MVN(0, Q) \\
&y_t = Zx_t + a + v_t \text{ where } v_t \sim MVN(0, R) \\
&x_0 \sim MVN(\Pi, \Lambda)
\end{aligned}
$$

We estimate the DFM using multivariate autoregressive state-space model[11]. The time-series of economic indicators (y) are modeled as linear combination of hidden trends (x) and factor loadings (Z) plus some offset 'a'. For example, the CEII-6 model consists of IIP-consumer goods, non-oil and non-gold imports, domestic auto sales, rail freight, air cargo and government receipts represented by six observed time series. It requires us to fit a model using a single-index dynamic factor, which we refer to as CEII. The MARSS specification consists of two stochastic components: an unobservable common component, x_t and an idiosyncratic component v_t. Both of these components are modelled as autoregressive stochastic processes. x_t is an estimate of CEII and Z represents the loadings of the economic indicators on the common component, CEII. The identifying assumption in the above model is that the co-movements in the time series indicators

[11] (MARSS) package in R

arise from the single source x_t, i.e. x_t enters each indicators with different loadings, Z_i, i=1…. ,6. This is ensured from our assumption that v_{it} and $x_t$are mutually uncorrelated at all leads and lags for all the 6 observed economic indicators. The same model is estimated further by including three additional variables (CEII-9) and six additional variables (CEII-12). It may be mentioned that all these three track actual GDP turning-points quite closely.

Next, we zoom into the dynamics CEII-6 in the recent quarters. Figure 2 includes both year-on-year (Y-o-Y) and month-on-month (M-o-M) variation in CEII-6. To iron out the short run fluctuations, we present a three-month moving average of both the series. The M-o-M series indicates monthly momentum in economic activities while the Y-o-Y series captures the yearly dynamics. The M-o-M series suggests sharp deceleration coinciding with the period of demonetization, but also a sharp recovery quickly thereafter. Subsequently, since early 2018, the monthly momentum suggests a gradual moderation in economic activity. The Y-o-Y series, like the monthly series, captures the demonetization downturn and the subsequent recovery, which peaks in December 2017, helped by the low base of the earlier year. The Y-o-Y series also reinforces the economic deceleration indicated in M-o-M series since early 2018 (Figure -2).

Figure-2: CEII-6, recent period dynamics

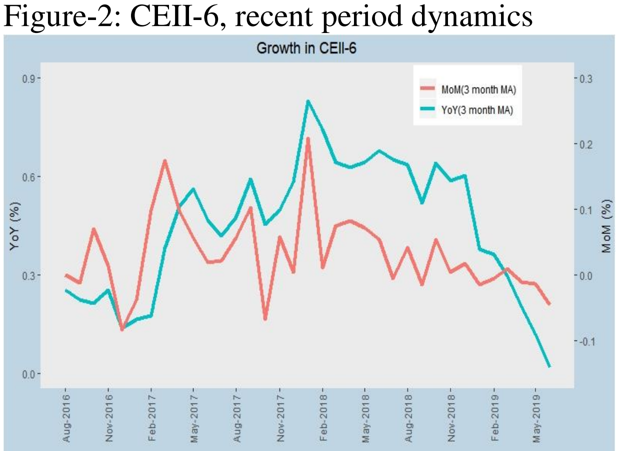
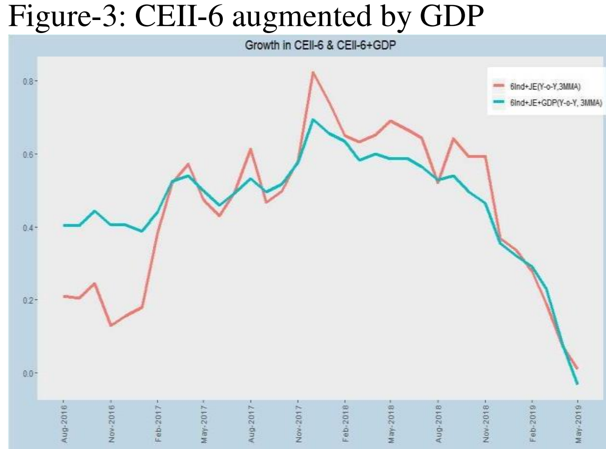


Figure-3: CEII-6 augmented by GDP



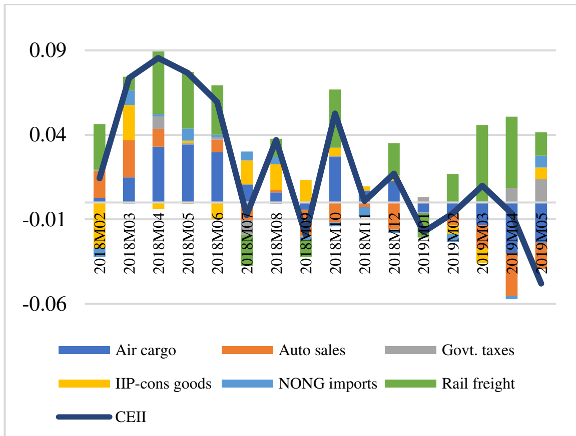
Some authors have emphasised including GDP along with other high frequency indicators of activity to construct coincident indicator (Dua et al. 1999). Following this literature, we also include GDP in CEII which is represented in Figure 3.[12] It may be mentioned that the dynamics of CEII doesn't change considerably with / without GDP and both indicate a downturn in economic activity in the recent period.

Figure-4 plots the contribution[13] of each of the indicators included in the model explaining in the variation in M-o-M CEII. We also report contributions of individual indicators in Table-2. It is evident that the recent deceleration has been mostly consumption driven. Auto sector and non-oil-non-gold imports present mixed picture, while rail freight has remained buoyant.

Table 2: Contributions by Components

	Jan-19	Feb-19	Mar-19	Apr-19	May-19
IIP-cons goods	0.002	-0.013	-0.016	0.029	
NONG imports	0.009	-0.016	0.006	0.003	0.012
Auto sales	0.033	0.003	-0.075	-0.003	0.031
Rail freight	-0.008	0.048	0.095	-0.016	-0.038
Air cargo	0.006	-0.043	-0.004	-0.043	
Govt. taxes	0.004	-0.003	0.001	0.027	
CEII	**0.046**	**-0.023**	**0.007**	**-0.003**	**-0.148**

Figure-4: Contributions by Components



[12] GDP is available until March 2019. The jagged edge methodology is used to handle the missing observations of GDP for recent months.
[13] Contribution of a component = M-o-M growth of component*Regression coeff. of M-o-M growth in component on M-o-M growth of CEII.

As indicated earlier, we expand our indicator list to include IIP-core, foreign tourist flows and exports. These additional variables represent investment and services activities. Figure-5 plots the recent dynamics of CEII-9, which is in line with CEII-6. However, both CEII-6 and CEII-9 underestimated (Figure-8) the actual decline in the quarter ending June 2017 as well as the recent slowdown. To address this problem in CEII-6 and CEII-9, we take recourse to including a financial block to improve the tracking of GDP growth, our target variable. While being fully aware of the fact that some of these variables might display leading properties and therefore deviate from the core principles of using only coincident indicators, we still included Sensex, bank credit and NEER to better track economic fluctuations (MSM No.10)[14]. CEII-12 points towards some recovery in economic activities, contrary to CEII-6 and CEII-9, in the recent times (Figure-6). It may be mentioned in thistext that bank credit and Sensex which were at their trough during the quarter ending June 2017, have been improving thereafter in recent times (Q4FY19). It may be mentioned that there are several other potential high-frequency financial variables (e.g. VIX, T-bill yields), which we intend to explore further.

Figure-5: CEII-9, recent period dynamics

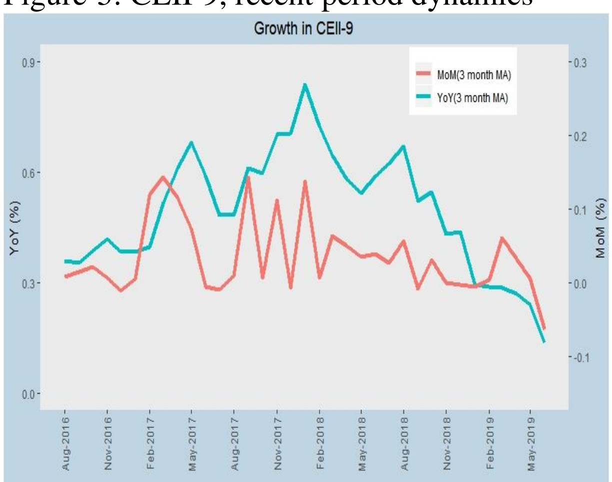
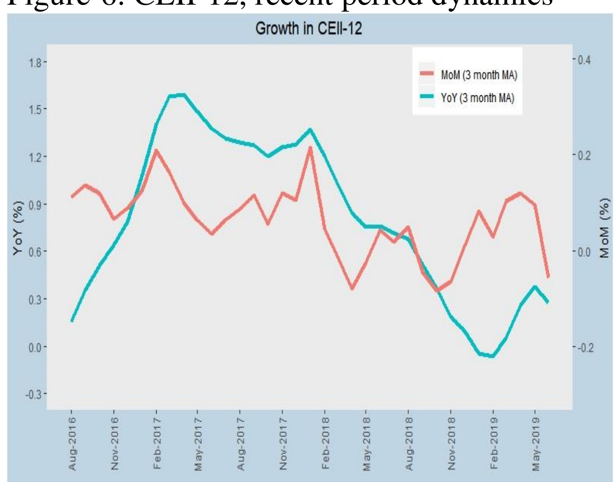


Figure-6: CEII-12, recent period dynamics



IV. 3. Nowcasting India's GDP Growth

Taking cue from the existing literature, we use the estimated CEIIs to nowcast current quarter GDP growth well before (at least 8 weeks in advance) the official release. For this purpose, we estimate a parsimonious AR model of GDP growth augmented by CEIIs (Y-o-Y). The nowcasts based on 6, 9 and 12 indicator models along with actual GDP growth are plotted in Figure 7. It is observed

[14] https://www.rbi.org.in/Scripts/MSM_Mintstreetmemos10.aspx

that the GDP-nowcasts track the actual GDP growth reasonably well over the estimation sample. Looking at the recent quarters, it is observed that while the nowcasts have tracked the turning points in GDP reasonably well, the nowcast based on CEII-6 and CEII-9 appear to have overestimated growth in the recent quarters.

Figure-7: GDP Growth and its Nowcasts

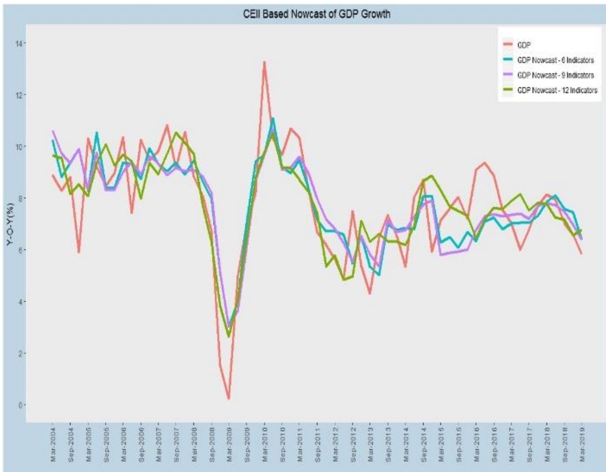
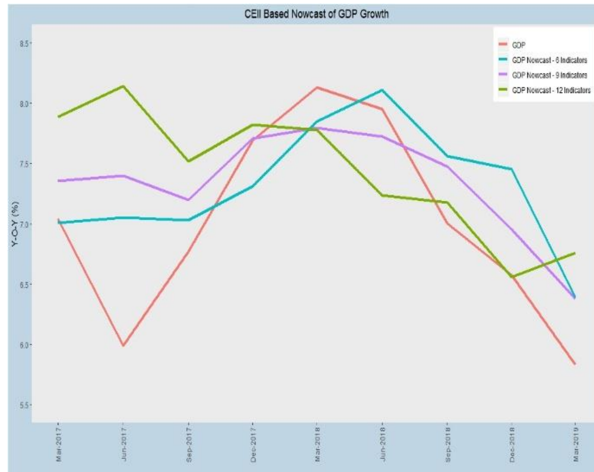


Figure-8: Recent Nowcasts



IV.4. Model Performance

The next step in model building pertains to evaluation of model performance. In this context, we examine the out-of-sample performance of the nowcasting models during 2017Q1-2019Q1. It is observed that the out-of-sample performance, measured in terms of root mean squared error (RMSE), is better for the CEII-6 model compared to the CEII-9 and CEII-12 models (Figure-9 and Table 3).

We are interested to know if there is any forecast accuracy gains obtained from using CEII estimates that have incremental information in the form of jagged edge data embedded in them. It may be mentioned here that hard data releases in India are non-synchronous in nature. For example, monthly production of coal and crude oil is typically released on the last working day of the month, monthly production of commercial vehicles is released during the middle of a month, and railway freight traffic of major commodities is released during the first 10 days of every month. Also, there are varying lags in data releases in India. Together, this results in jagged-edge data. The out-of-sample performances suggest that CEII estimates obtained from jagged-edge data structure perform marginally better than those CEII estimates obtained from data set without rough edges. The jagged-edge CEII estimates consistently record a lower RMSE value for 6, 9 and 12-indicator models compared to its CEII counterparts that are without jagged edges.

Next, we apply the mixed data sampling (MIDAS) regression to exploit the rich information contained in our monthly CEII to better nowcast quarterly GDP and improve out-of-sample performance. We compare between models that adopt the regular OLS method versus those that adopt the MIDAS method, which are specially equipped to handle mixed frequency data. In general, the baseline CEII-6 model performs better out-of-sample compared to the CEII-9 and CEII-12 models. However, we do not observe much of a forecast accuracy gains from applying MIDAS over OLS regression for the CEII-6, CEII-9 and CEII-12.

Figure-9: Out of Sample performance

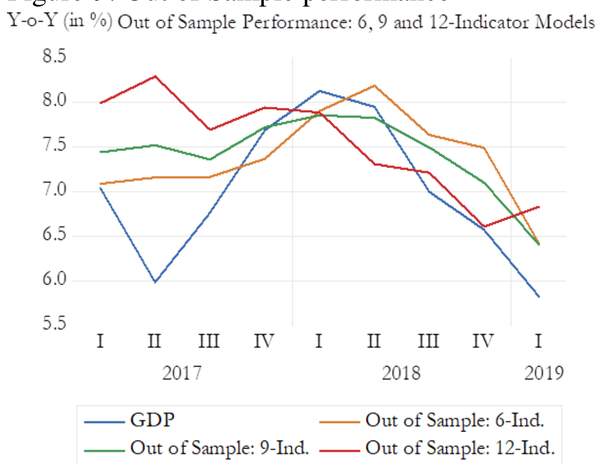


Table-3: Root Mean Squared Errors

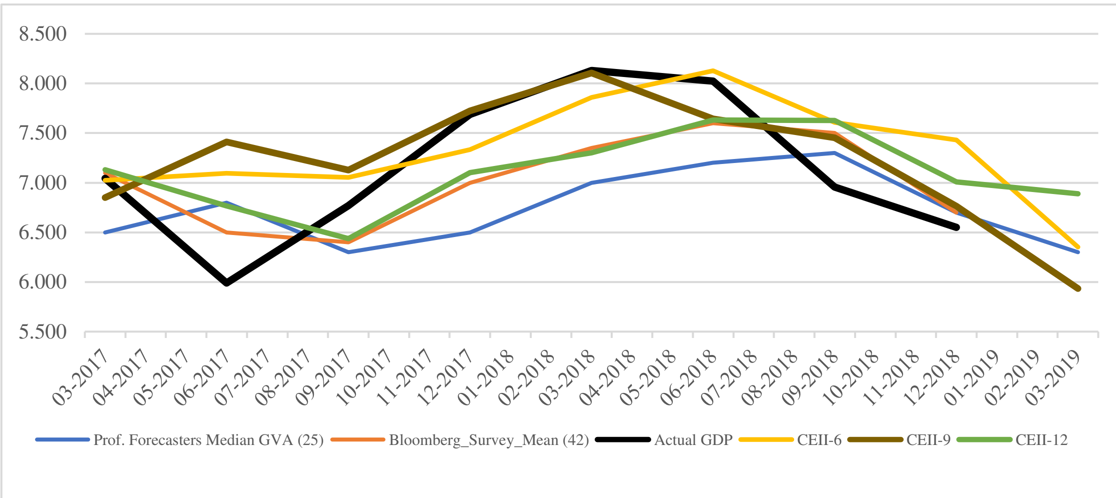
Out of Sample (2017Q1-2019Q1) Performance: Root Mean Squared Error

	MIDAS	Jagged Edge: Yes	Jagged Edge: No
6-Indicator	Yes	0.721	0.718
	No	0.609	0.616
9-Indicator	Yes	0.843	0.842
	No	0.651	0.654
12-Indicator	Yes	1.067	1.038
	No	0.980	0.982

As a comparison of the different GDP nowcast performance across models, we report the out-of-sample performance of 6-indicator, 9-indicator and 12-indicator models. Table-3 reports the out-of-sample root mean squared error (RMSE) value of 9-quarter ahead nowcasts i.e. 2017 Q1-2019 Q1. We attempt comparing between models that utilize dataset with jagged edges and use MIDAS vis-à-vis those that do not, for the 6-indicator, 9-indicator and 12-indicator models. The 6-indicator model using the jagged edge dataset records the lowest out-of-sample RMSE. Broadly, the analysis from using the jagged edge dataset seems to suggest that by exploiting the incremental information that is available from the actual flow of data releases, we are able to improve model performance.

We evaluate the performance of CEII with other market projections of GDP growth. In this context, as mentioned in Section II, professional economists use their proprietary models, and they share their projections with Bloomberg and RBI (professional forecasters' survey). We plot the median (mean) of such projections along with CEII nowcasts (Figure-10). Among the nowcast models that are available in the public domain, Rabobank suggests GDP growth nowcast at 6.3% while Bloomberg's monthly GDP Growth tracker projects the same at 6.5% in Q4 of FY 19. It may be observed that CEII-based nowcasts capture the turning points and closely tracks GDP compared with other market forecasts.

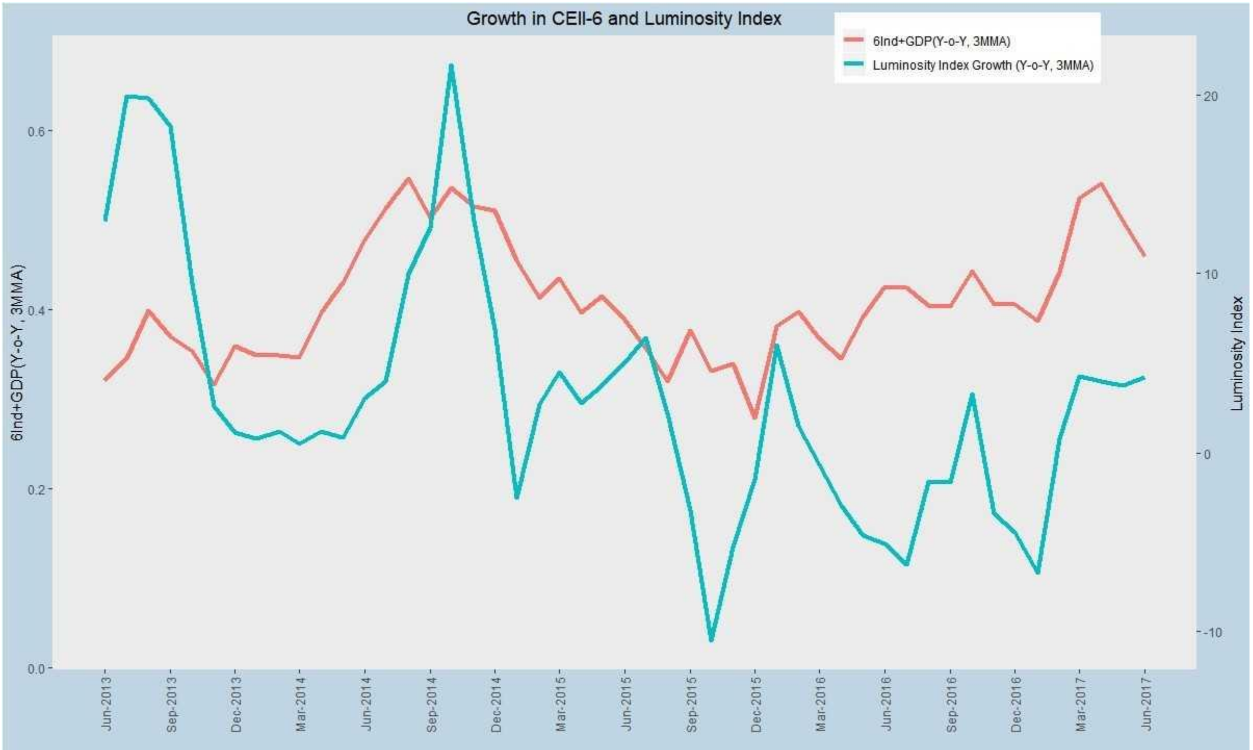
Figure-10: Nowcast Performance Evaluation vis-a-vis Market Projections



Finally, as a test of robustness, we compare the dynamics of CEII with an alternative and unconventional measure of economic activity captured in the nightlight data (luminosity index).[15] Notwithstanding the seasonality (September–October spike), limited availability (upto 2017) and aggregation problems (national average), luminosity index is expected to capture the economic activities well, particularly in economies with dominant informal sectors. It is reassuring to note that Y-o-Y dynamics of CEII corresponds with the nightlight data, which is provided by an independent source and is being extensively used of late (Figure 11).

[15] Infrared Imaging Radiometer Suite (VIIRS) of National Aeronautics and Space Administration (NASA) and National Oceanic and Atmospheric Administration (NOAA), USA. VIIRS data has a wider radiometric detection range than former generation of similar satellites, which solves the issue of over-saturation at bright core centres (Elvidge et al. 2013). However, the publicly available VIIRS data still requires processing before use, as some temporary lights and background noise remain. We follow the procedure discussed in Beyer et al. (2018) and remove all observations from areas categorized as background noise mask. After outlier removal, these areas are identified by clustering the remaining observations based on their intensity.

Figure-11: Growth in CEII-6 and Luminosity Index



V. *Conclusion*

A core concern in policymaking is identifying the signs of expansions and contractions in economic activity. At any point in time, diverse economic activity indicators may indicate mixed trends. Therefore, combining all of these together in an appropriate way to arrive at the underlying (or unobserved) trend has traditionally occupied the attention of both governments and businesses. Our aim has been to contribute to the existing literature by combining high frequency indicators, which are useful proxies of economic activity, to nowcast GDP growth of India.

To provide an early estimate of the current quarter GDP growth, we construct a Coincident Economic Indicators for India (CEII) using 6, 9 and 12 high-frequency indicators. These indicators represent various sectors, display high contemporaneous correlation with GDP, and co-move in line with the GDP turning points. While CEII-6 includes domestic economic activity indicators, CEII-9 combines indicators on trade and services along with the indicators used in CEII-6. Finally, CEII-12 adds financial indicators to the indicators used in CEII-9. In addition to the conventional economic activity indicators, we include a financial block in CEII-12 to reflect the growing influence of the financial sector on economic activity.

CEII is estimated using a dynamic factor model to extract a common trend underlying the high-frequency indicators. We use the underlying trend to gauge the state of the economy and to identify sectors contributing to economic fluctuations. Further, CEIIs are used to nowcast GDP growth, which closely tracks the actual GDP growth, both in-sample and out-of-sample.

VI. References

Bańbura, M., & Modugno, M. (2014). “Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data”. Journal of Applied Econometrics, 29(1), 133-160.

Bańbura, M., Giannone, D. and L. Reichlin (2010). “Large Bayesian vector auto regressions”. Journal of Applied Econometrics, 25(1), 71-92.

Banbura, Marta and Giannone, Domenico and Reichlin, Lucrezia, (2010) “Nowcasting”. European Central Bank Working Paper No. 1275: https://ssrn.com/abstract=1717887

Bhadury S & S Pohit & R. Beyer, (2018). "A New Approach to Nowcasting Indian Gross Value Added," NCAER Working Papers 115, National Council of Applied Economic Research.

Bhattacharya, R., Pandey, R., & Veronese, G. (2011). “Tracking India growth in real time” National Institute of Public Finance and Policy. (pp. 2011-90).

Bok, Brandyn, et al. (2018) "Macroeconomic nowcasting and forecasting with big data." Annual Review of Economics 10: 615-643.

Bragoli, D., & Fosten, J. (2018). “Nowcasting Indian GDP”. Oxford Bulletin of Economics and Statistics, 80(2), 259-282.

Burns, A. F., & Mitchell, W. C. (1946). “The basic measures of cyclical behavior in Measuring Business Cycles”. National Bureau of Economic Research. (pp. 115-202).

Caruso, Alberto. (2015) "Nowcasting Mexican GDP." European Center for Advanced Research in Economics and Statistics working paper No. 40

Kyosuke Chikamatsu, Naohisa Hirakata, Yosuke Kido, Kazuki Otaka, 2018. "Nowcasting Japanese GDPs," Bank of Japan, Bank of Japan Working Paper Series 18-E-18,

Dahlhaus, T., Guénette, J. D., & Vasishtha, G. (2017). “Nowcasting BRIC+ M in real time”. International Journal of Forecasting, 33(4), 915-935.

Dua, P., & Banerji, A. (2000). “An indicator approach to business and growth rate cycles: The case of India”. Indian Economic Review, 55-78.

Dua, P., & Sharma, V. (2016). “A comparison of economic indicator analysis and Markov switching methods concerning the cycle phase dynamics”. OECD Journal: Journal of Business Cycle Measurement and Analysis, 2015(2), 1-27.

Ghysels, Eric, Santa-Clara, Pedro and Valkanov, Rossen, (2004), The MIDAS Touch: Mixed Data Sampling Regression Models, CIRANO Working Papers, CIRANO, https://EconPapers.repec.org/RePEc:cir:cirwor:2004s-20.

Giannone, D., Reichlin, L., & Small, D. (2008). “Nowcasting: The real-time informational content of macroeconomic data”. Journal of Monetary Economics, 55(4), 665-676.

Kumar, Gitanjali (2013) : High-frequency real economic activity indicator for Canada, Bank of Canada Working Paper, No. 2013-42, Bank of Canada, Ottawa

Luciani, M., Pundit, M., Ramayandi, A., & Veronese, G. (2018). "Nowcasting Indonesia". Empirical Economics, 55(2), 597-619.

RBI (2002). "Report of the Technical Advisory Group on Development of Leading Economic Indicators for Indian Economy".

RBI(2007). "Report of the Technical Advisory Group on Development of Leading Economic Indicators for Indian Economy."

Stock, J. H., & Watson, M. W. (1989). "New indexes of coincident and leading economic indicators". NBER Macroeconomics Annual, 4, 351-394.

Annex Table I: International literature on Nowcasting GDP

Institutions	Bank of England	Federal Reserve Bank of Atlanta (FRBA)	Federal Reserve Bank of New York (FRBNY)	European Central Bank	Norges Bank	Canada	Japan	Indonesia	Mexico
Blocs									
Industry and construction	Total Business Investment Housing Investment Construction Output Industrial Production Manufacturing Production PMI-Construction PMI-Manufacturing CIPS-E-Manufacturing CIPS-E-Construction CBI Industrial Trends CBI Distributive Trends Lloyds Business Barometer Agents' Score	Philly Fed Business Outlook Survey Chicago Fed Midwest Manufacturing Index ISM Manufacturing Index Industrial Production New Residential Construction C20 (Housing Starts) Manufacturers' Shipments, Inventories and Orders Construction Spending	Housing Starts Building Permits Value of Construction put in place Industrial Production Index Capacity Utilization Inventories: Total Business Merchant Wholesalers Inventories: Total Manufacturers' New Orders: Durable Goods Manufacturers' Shipments:Durable Goods Manufacturers' Inventories: Durable Goods Manufacturers' Unifilled Orders: All Manufacturing Industries Merchant Wholesalers Inventories: Total Inventories: Total Business Empire State Mfg. Survey: General Business Condition Phila.Fed.Mfg.Business Outlook: Current Activity ISM mfg.: PMI Coomposite Index ISM mfg.: Price Index	Industrial Production: Total Industry Industrial Production: Manufacturing New Orders: Manufacturing Working on Orders PMI-Manufacturing Industrial Confidence Indicator Retail Trade Confidence Indicator	PMI Industrial Production Turnover: Mining/Manufacturing Construction Output Capacity Utilisation Industrial Confidence Indicator New Orders: All Industries	Manufacturing Shipments	IIP Index of Tertiary Industry Activity Index of All Industry Activity Index of Construction Industry Activity Current Survey of Commerce (Sales Value, Wholesale) Survey of Production Forecast	PMI-Manufacturing	IMEF Business Climate Index-Manufacturing IMEF Business Climate Index-Non-manufacturing Producer Confidence Index Opinion Survey-Manufacturing Orders Total Vehicle Production Industrial Production Production of Crude Producer Confidence Index
Personal Income/ Consumption	Retail Sales Index Private Consumption	Reuters/University of Michigan Index Conference Board Consumer Confidence BEA Unit Auto Sales Existing Home Sales New Home Sales Perosonal Income and Sales	Real Personal Consumption Expenditure Real Disposable Personal Income New Single Family Houses Sold Retail Sales and Food Services	New Passenger Car Registration Consumer Confidence Indicator	Retail Sales Consumer Confidence	Retail Sales		Consumer Confidence Index Danareksa Consumer Confidence Local Auto Sales Motorcycle Sales	Consumer Confidence Index Automobile Sales Truck Sales Retail Sales
Employment	Claimant Count Rate LFS Number of Employees LFS Unemployment Rate	Initial Unemployment Insurance Claims Employment Situation	JOLTS Job Openings: Total ADP Nonfarm Private Payroll Employment Nonfarm Business Sector: Unit Labour Cost Civilian Unemployment Rate All Employees: Total Nonfarm	Unemployment Rate: Total Index of Employment: Total Industry	Employment Unemployment Rate	Monthly Payroll Employment			Unemployment Rate
Services	Index of Services CPIS-E-Services PMI-Services			Purchasing Managers Survey, Services: Business Activity Services Confidence Indicator					
External Sector	Sterling Effective Exchange Rate BOP Total Imports BOP Total Exports	International Trade	Exports: Goods and Services Imports: Goods and Services	Extra Euro Area Trade: Export Value Nominal Effective Exchange Rate: Core Group of Currencies against Euro	Merchandise Exports Merchandise Imports			Foreign Reserve Trade Balance Exports Net Foreign Assets Imports BoP Current Account Balance	Total Vehicle Exports Crude Exports Crude Imports Trade Balance
Prices		U.S. Imports Price Index U.S. Exports Price Index Producer Price Index Consumer Price Index	PCE: Chain Price Index PCE less Food and Energy: Chain Price Index CPI-U: All Items CPI-U: All Items less Food and Energy	HICP: Overall Index PPI: Excluding Construction Consumer Survey: Price Trend (12 months) Industry Survey: Selling price Expectation Raw Material, excl. Energy: Market Prices Raw Material, Crude Oil: Market Prices	CPI		PPI (deflator for Sales Value)	CPI CPI-Core CPI-NSA	
Miscellaneous Economic Activity	GDP	ISM Non-manufacturing Index Wholesale Trade Retail trade and inventories Manufactured Home Surveys	ISM Non-manufacturing: NMI Composite Index Real GDP	GDP: Chain Linked	GDP Mainland Norway	GDP	GDP Reuters Tankan DI (soft indicators) Economy Watchers Survey	GDP (YoY and QoQ)	
Credit and Finance	Mortgages Approved Net Consumer Credit UK Focussed Equity Index Term Spread Corporate Bond Spread	S&P 500 Index		M3: Index of National Stocks Index of Loans Dow Jones Euro Stoxx: Broad Stock Exchange Index Euribor 3 month				Bank Indonesia Reference Rate Money Supply (M1) Money Supply (M2)	
Methodology									
	Release-Augmented Dynamic Factor Model	Dynamic Factor Model Bridge Equation Approach	Dynamic Factor Model Kalman-Filtering Techniques	Dynamic Factor Model Kalman-Filtering Techniques	Bayesian Dynamic Factor Model	Dynamic Factor Model Kalman-Filtering Techniques	Dynamic Factor Model Mixed Data Sampling	Dynamic Factor Model	Dynamic Factor Model Expectation- Maximization Algorithm

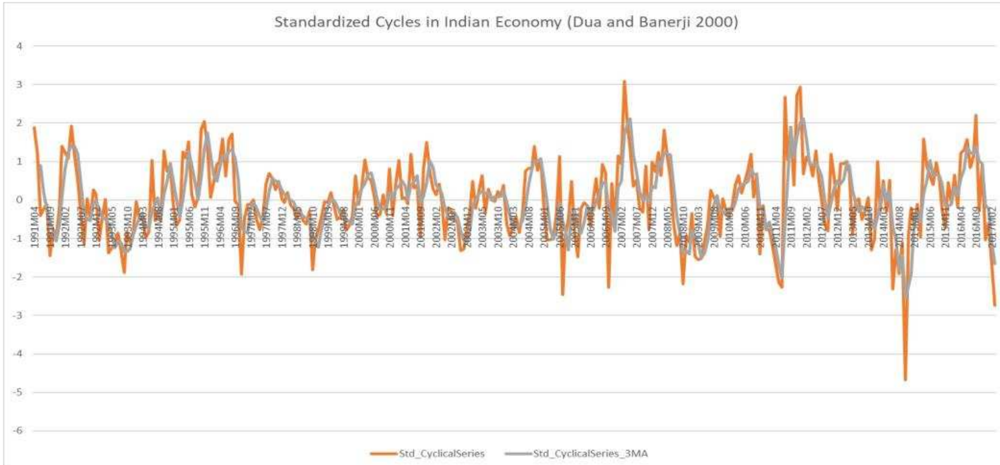
Annex Table II: Nowcasting literature for India

Institutions	RBI(CEII-6)	RBI(CEII-9)	RBI(CEII-12)	Bloomberg	Rabobank	NIPFP (2011)	Dalhaus et al (2017)	Bragoli and Fosten (2016)	Dua and Banerjee (1999)	RBI (2006)
Blocs				Regularly Updated		One Time Study				Group Report
Industry and construction		IIP-Infrastructure	IIP-Infrastructure	Industrial Production Capital Expenditure Index Construction Index	PMI-Manufacturing Electricity Production IIP Crude and Steel Production	IIP-Mining IIP-Manufacturing IIP-Electricity Production of Coal and Crude Production of Cement and Steel Commercial Vehicle Production	PMI-Manufacturing IIP-Basic Metal IIP-Electricity	IIP PMI-Manufacturing Production of Crude Production of Steel Electricity Generation Industrial Performance Assessment	Industrial production of consumer goods	IIP-General Index IIP-Basic Metal IIP-Electricity IIP-Intermediate Goods Production of Commercial Motor Vehcile Cargo Handled in Major Ports
Personal Income/ Consumption	IIP-Consumer Goods Auto Sales	IIP-Consumer Goods Auto Sales Petroleum Product Consumption	IIP-Consumer Goods Auto Sales Petroleum Product Consumption	Auto Sales Petroleum Product Consumption	Vehicle Sales Car Passenger Registration Personal Loans Petroleum Product Consumption	Cellular Subscription				
Employment									Monthly Registered Unemployed	
Services	Government Tax Revenue	Government Tax Revenue	Government Tax Revenue		PMI-Services Net Tax Revenue	Central Government Revenue Expenditure				
External Sector	Non-Oil and Non-Gold Import	Non-Oil and Non-Gold Import	Non-Oil and Non-Gold Import	Agricultural Trade Balance	Vehicle Export		Exports Imports REER	Exchange Rate (INR/USD)		Exports Non-Oil Imports
Prices					CPI	WPI	CPI WPI Petrol Spot Price (Brent) World Commdity Price Index	CPI-Industrial Worker CPI- Agricultural Labourers CPI-Rural Labourers WPI-All Items		WPI: Industrial Raw Materials WPI: Manufactured Products Gold Prices in Mumbai
Miscellaneous Economic Activity	Rail Freight Air Cargo	Rail Freight Air Cargo Foreign Tourist Arrival	Rail Freight Air Cargo Foreign Tourist Arrival	Foreign Tourist Arrival Traffic Index	Volatility Index Railway Freight Earnings	Railway Goods Traffic Port Traffic	ISM Composite Index S&P 500 Composite Index	US Industrial Production US ISM PMI-Manufacturing Euro Area 19 Industrial Production Euro Zone PMI-Manufacturing Asia Sentix Overall Index	GDP at factor cost interpolated to a monthly series Wages to worker in factory sector	US GDP US Leading Indicator Index Euro Area Leading Indicator Index
Credit and Finance				Real Currency Demand Real Combined Credit	Industry Loans Services Sector Loans MIBOR BSE-500 BSE-Sensitive Index Monetary Base (M0)	Non-Food Bank Credit Deposits NSE Turnover	Money Supply NSE-500 91-day Tbill 10 Year Bond Yield FOMC-Fed Fund Target Rate	Money Supply (M1) 91-day Tbill Sensex BSE 30		Broad Money (M3) Real M3 (M3/WPI) Currency with Public Bank Credit Forward Premia 6-month
Methodology	Dynamic Factor model Mixed Data Sampling	Dynamic Factor Model Mixed Data Sampling	Dynamic Factor Model Mixed Data Sampling	Weighted Average of Monthly Activity Indicators	Bayesian VAR Ordinary least squares Combined model	Bridge Equation Models	Dynamic Factor Model	Dynamic Factor Model		Dynamic Factor Analysis Turning Point Analysis Cross-Correlation Analysis

Annex Table III: Dynamic cross-correlation for indicator selection

	GDP (t-3)	GDP (t-2)	GDP (t-1)	GDP (t)	GDP (t+1)	GDP (t+2)	GDP (t+3)	Indicator Type
AGRI WAGES	-0.01	-0.21	-0.23	0.01	-0.07	-0.11	0.01	X
AIR CARGO	0.16	0.01	0.29*	0.46*	0.23	-0.13	-0.12	C
AIR PASSENGER	0.06	0.13	0.35*	0.35*	0.33*	0.14	-0.09	C
AUTO PASSENGER	-0.11	0.01	0.30*	0.23	0.29*	0.04	-0.33*	L-/L+
AUTO COMMERCIAL	0.08	-0.21	0.22	0.50*	0.21	0.06	0.01	C
AUTO TOTAL	0.05	-0.09	0.13	0.30*	0.24	-0.06	-0.14	C
BANK CREDIT	-0.09	0.02	-0.06	0.02	0.09	0.23	-0.02	X
CEMENT	0.02	-0.02	0.13	0.07	0.04	-0.32*	-0.11	X/L+
CPIIW	0.00	-0.11	-0.09	-0.02	0.06	-0.07	-0.23	X
CRBCOMM	-0.08	-0.32*	-0.04	0.44*	0.51*	0.34*	0.17	C/L+
CRUDE INDIAN BASKET	-0.004	-0.37	-0.21	0.32*	0.37*	0.13	0.10	C/L+
EXPORTS	-0.11	-0.31*	-0.08	0.30*	0.41*	0.33*	0.15	C/L+
FOREIGN TOURIST	0.17	-0.07	0.00	0.34*	0.23	0.04	0.03	C
FOREX	-0.23	-0.20	0.20	0.36*	0.36*	0.39*	0.15	C/L+
GOVT. RECEIPT	0.27*	-0.03	0.02	-0.04	0.25*	-0.08	-0.07	L+
IIP CONSUMER	0.07	0.04	-0.15	0.39*	0.52*	-0.07	-0.09	C/L+
IIP INFRA	0.26*	-0.26*	0.09	0.30*	0.28*	-0.08	-0.03	C/L+
NEER	0.09	0.13	0.23	0.26*	0.20	0.12	0.17	C
NONG	0.04	-0.07	-0.11	0.02	0.32*	0.31*	0.30*	L+
OIL CONSUMPTION	0.05	-0.05	-0.01	0.17	0.09	-0.16	-0.01	X
PERSONAL LOANS	-0.03	0.11	-0.13	0.10	0.33*	0.20	0.25	L+
RAIL FREIGHT	-0.07	-0.12	0.13	0.28*	0.23	-0.31*	-0.15	C
RAIL PASSENGER	0.01	0.00	0.00	0.12	0.03	-0.04	0.14	X
SENSEX	0.02	-0.05	0.23	0.54*	0.41*	0.06	-0.10	C
STEEL	-0.04	-0.01	0.05	0.12	0.28*	0.02	0.02	L+
VIX	-0.0042	0.24	-0.04	-0.36*	-0.21	0.11	0.13	C
T BILL91	-0.11	-0.25*	-0.36*	-0.33*	-0.25	-0.15	-0.004	C/L-

Annex Note I: Replication of Dua and Banerji (2000)[16]



Step involved in replication and updating

1. All data monthly, and seasonally (Census X-11) adjusted
    a. We use Census X-12 instead. We also obtain the cyclical component of the log deviation of each series by applying an H-P Filter ($\lambda = 14400$). We then standardize each cyclical series using their respective standard deviations.
2. Output: Two different measures of CEI which correspond to the following two different variables for output:
    a. <u>Real GDP at Factor cost</u>: Das (1993) has quarterly data on GDP for the period 1970-91. This is interpolated to 1950. Monthly data - dividing quarterly data by 3.
    b. Closely follow Banerji and Dua (2000), and obtain monthly GDP series using the procedure given in Das (1993). For data from 2012-13, we use the Real GVA at FC
3. <u>Index of Industrial Production</u>: Monthly data available
    a. We do not report results using IIP for our exercise (also not reported in Banerji and Dua (2000)).
4. Income: Annual wage data from the ASI are interpolated into monthly data using a monthly adjustment factor.
    a. <u>The Adjustment Factor</u>: the adjustment factor = [monthly variation in the consumer manufacturing output relative to an annual average] X [relative volatility of annual wages and annualized consumer manufacturing output]
    b. The closest data available is annual compensation of employees, from 1991 from the EPW-Research foundation. This is collated from ASI.

Employment: Monthly seasonally adjusted unemployment numbers from the Monthly Abstract of Statistics (MAS)

    c. This data is available only until Dec-2013. MAS is no longer published. One option is to collate all the previous MAS and forecast present data using suitable forecast methodology.
    d. We use the total labor force data and multiply this with the annual unemployment rate – available with world bank from 1991 – to obtain the annual unemployment data. We then follow the same interpolation used by Banerji and Dua (2000) to obtain the monthly unemployment series.

5. Trade: In this paper, trade is all domestic. Industrial production of consumer goods at constant prices and seasonally adjusted
6. Data on actual production levels are not available. We therefore use monthly consumer goods IIP – after necessary splicing to get monthly series at the 1980-81 base.

[16] SRU Mimeo, Gopalakrishnan, Kumar and Ghosh (2017)

Annex Note I: *Jagged Edge Data*

Principal components come from the largest Eigenvalues of the sample correlation matrix of the series,

$$S = \tfrac{1}{T}\sum_{i=1}^{T} X_t X_t^{'}$$

The r largest principal components are extracted from the sample correlation matrix. D is the $r \times r$ diagonal matrix with diagonal elements given by the largest r Eigenvalues of S, and denoted by V the $n \times r$ matrix corresponding Eigenvectors s.t the normalisation gives $V^{'}V = I_r$. Following is the approximation of the common factors:

$$\tilde{F} = V^{'} X_t$$

Once we have estimated the common factors, $\tilde{F}$, we can estimate the factor loadings, Γ, and the covariance matrix of the idiosyncratic components, Π. This is done by regressing the data series on the estimated common factors, as follows:

$$\hat{\Gamma} = \sum_t X_t \tilde{F}_t^{'} (\tilde{F}_t \tilde{F}_t^{'})^{-1} = V$$

The estimated covariance matrix of the idiosyncratic components, $\hat{\Pi}$, is as follows:

$$\hat{\Pi} = diag(S - VDV)$$

The dynamic factor equation parameters, A and B, can be estimated from VAR, on the common factors, $\tilde{F}_t$, where $F_t = AF_{t-1} + Bu_t$. These estimates, $\hat{\Gamma}, \hat{\Pi}, \hat{A}, \hat{B}$, have been proven to be consistent as $n, T \to \infty$ by Forni et al. (2000). Given the estimated parameters, in the second step, an updated estimate of the common factors is obtained using the Kalman smoother.

Annex Note II: MIDAS

The explanatory variables can also have frequencies different from each other. The basic equation for MIDAS is similar to that of distributed lag models, exhibiting a dynamic relationship between the dependent and independent variables. However, there are still significant differences between the two methods. The basic equation for MIDAS is (Ghysels et al. 2004):

$$Y_t = \beta_0 + B\left(L^{\frac{1}{m}}\right) X_t^{(m)} + \epsilon_t^{(m)}$$

Where $B\left(L^{\frac{1}{m}}\right) = \sum_{j=0}^{j^{max}} B(j) L^{j/m}$ is a polynomial of length j^{max} in the $L^{1/m}$ operator and the $L^{j/m}$ operator lags $X_t^{(m)}$ by j/m periods. We can run the above non-linear regression to estimate the dependent variable. Extracting the maximum information from the dataset, requires a suitable polynomial and a suitable polynomial may involve an increased number of lags of $X_{(t-j)}^{m}$ data. This requires estimating many parameters and is one of the shortcomings of the MIDAS (Ghysels et al. 2004).