Tracking India Growth in Real Time

Rudrani Bhattacharya, Radhika Pandey and Giovanni Veronese

Working Paper No. 2011-90

July 2011

National Institute of Public Finance and Policy
New Delhi
http://www.nipfp.org.in

Tracking India Growth in Real Time

Rudrani Bhattacharya
Radhika Pandey
Giovanni Veronese [*]

Abstract

Tracking growth in the Indian economy would be best performed using a measure like GDP. Unfortunately official estimates of this indicator are released with quarterly frequency and with considerable delay. This paper compares different approaches to the short term forecasting (nowcasting) of real GDP growth in India and evaluates methods to optimally gauge the current state of the economy. Univariate quarterly models are compared with bridge models that exploit the available monthly indicators containing information on current quarter developments. In the forecasting exercise we perform a pseudo real-time simulation: by properly taking into account the actual publication lags of the series, we replicate the information set available to the policymaker at each point of time. We find that bridge models perform satisfactorily in predicting current quarter GDP growth. This result follows from the actual estimation technique used to construct the official quarterly national accounts, still largely dependent on a narrow information set. Our analysis also suggests mixed evidences about the additional predictive power of Indian survey data with respect to the hard data already used in the national accounts.

JEL classification: C22; C32; C53 Keywords: Nowcasting; Bridge model; Factor model; Emerging markets; India

---

[*] Rudrani Bhattacharya and Radhika Pandey are with the National Institute of Public Finance and Policy and Giovanni Veronese is with the Bank of Italy. The authors gratefully acknowledge encouragement and support from Ajay Shah, Ila Patnaik, and the team at National Institute of Public Finance and Policy. The views expressed in this paper are exclusively those of the authors and not of their institutions. This paper was written under the aegis of the NIPFP-DEA Research Program.

Tracking India Growth in Real Time

Introduction

Access to timely and reliable information on the current state of economic activity is essential for effective policy making. Correct initial conditions are crucial ingredients for meaningful forecasting exercises, often conducted on the basis of large structural models, which are required to support a forward looking policy framework.

To obtain these early estimates, or nowcasts, economists resort to information from data which are related to the target variable (GDP or some subcomponent) but that are collected at higher frequency (monthly, weekly, daily) and released in a more timely manner. The academic literature on nowcasting methods has expanded rapidly in the last decade. Building from the simple bridge equations based on a narrow set of indicators (Baffigi *et al.*, 2003) the modelling has become increasingly complex to account for the larger information sets available nowadays, and to properly formalise the process of the information updating that occurs when data become available, or are revised. This is typically done relying on large state space factor models techniques (Giannone *et al.*, 2008).

While most of the nowcasting literature focused on developed economies where high frequency data on a large number of variables are published in timely manner (Barhoumi *et al.*, 2008; Giannone *et al.*, 2009; Banbura and Giannone, 2010; Angelini *et al.*, 2008; Knut and Trovik, 2007; D'Agostino *et al.*, 2008), nowcasting of economic activities in emerging markets with informational scarcity is rarely attempted. Several challenges emerge in the process of nowcasting of economic activity in emerging markets such as poor quality of data, short sample period for which indicators are available and the possibility of structural break in the economic time series that affects the choice of appropriate model (Maier, 2011). To our knowledge, (Pedersen, 2010; Maier, 2011; Matheson, 2011) are the very few attempts of nowcasting economic activities in emerging markets. Pedersen (2010) nowcasts Chilean GDP by extracting signals from monthly indicators. Maier (2011) evaluates different approaches for using monthly indicators to nowcast and forecast Chinese GDP, while Matheson (2011) predicts economic activity of a large number of countries, including emerging markets at a monthly frequency by utilising a wide range of economic time series in a timely fashion.

India has emerged as one of the important players in the world economy over the past two decades. It is one of the fast growing emerging economies, rapidly integrating with the world economy over this period, accounting for 2 percent of the value of world output at nominal exchange rates (almost 6% in US$ on a PPP basis) (Winters and Yusuf, 2007), and is one of the major players in world FDI flows. Given India's increasing role in the global economy, and its increased synchronisation with the global business cycle, timely prediction of the pulse of the Indian economy is important to gauge the dynamics of world output (Borin *et al.*, 2010). To this end, we attempt to nowcast the Indian GDP growth and evaluate, for the first time in the literature, the information content of the monthly indicators available in India, a country characterized by still a very scant set of high frequency statistics.

Official estimates of GDP in India are released with considerable delay, suffer from sizeable revisions and are not available in seasonally adjusted format. The first release of quarterly GDP growth is published approximately seven to eight weeks after the end of the reference quarter. This delay leads most analysts to look elsewhere to form their views, considering disparate indicators available at a higher frequency, which provide only a partial representation of overall economic activity and may contain significant idiosyncratic noise. In particular, business surveys which result from a qualitative assessment recorded in firms's interviews are only indirectly related to growth dynamics in the official measures of GDP.

In contrast, other monthly indicators are themselves part of the inputs into the quarterly national account computations performed by the Central Statistical Organisation (CSO). However, the estimation of GDP is generally complex and difficult to replicate, as the statistical institute may have access to additional sources, not available to the public, and because the exact estimation methodology remains confidential. What we know, is that in India, the reference figures for quarterly GDP are computed from the production side, aggregating estimates of the Value Added in each sector of the economy, which rely on various proxy indicators of economic activity.

In this paper we evaluate alternative methods that exploit timely monthly releases to compute early estimates of current quarter national accounts aggregates. The evaluation is conducted using an out of sample forecasting exercise. Namely, we perform a pseudo real-time simulation: by properly taking into account the actual publication lags of the various monthly series, we replicate the information set available to the policy maker at each point in time, and nowcast the upcoming GDP data release.

We restrict our nowcasting analysis to two measures of growth, based on GDP — excluding agriculture (GDPXagri) and GDP excluding agriculture and other services (GDPXoth). Our choice is motivated by the fact that some of these sectors' developments are unrelated to the business cycle movements of the economy, and display considerable volatility. In particular, Indian agriculture is still affected by strong seasonal oscillations which depend on the outcome of the monsoon. In contrast, the sector "other services", mainly composed of government services is affected by significant short run volatility due to the dynamics of public sector outlays.

Our findings show that an effective nowcast of GDP in the Indian context can be performed by using a multi-sectoral bridge model that strives to mimic as closely as possible the national accounts estimation procedure.

We find that bridge models relying on a small set of pre-selected key monthly indicators, serving as proxies for the various sub-sectors of the economy perform satisfactorily in predicting current quarter GDPXagri and GDPXoth growth. The performance of these models is compared with the benchmark quarterly auto-regressive and naive models. We find that the bridge models significantly outperform these benchmarks. The multi-sectoral bridge model also outperforms simple bridge models relying on single indicators (e.g. industrial production, global survey data). We also find that large state space factor models fall behind the bridge model in terms of forecast evaluation in pseudo real-time. Matheson (2011) also finds factor model to perform poorly for India along with Australia and Saudi Arabia.

Our results also provide substantial evidence that the actual estimation technique used by the Central Statistical Organization (CSO) to construct the official quarterly national accounts, is still largely dependent on a rather narrow information set.

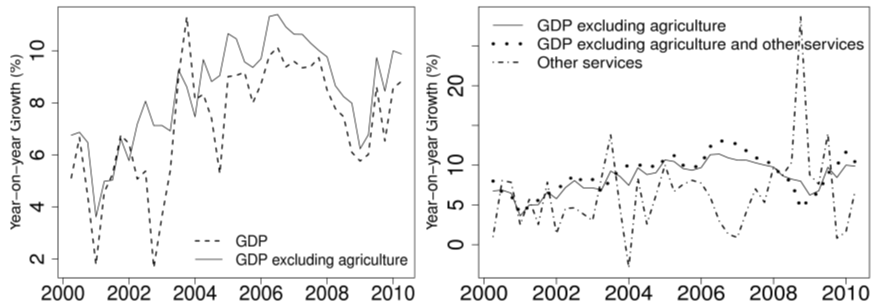
Finally, we investigate for the first time the effective usefulness of Indian survey data in nowcasting GDP. The literature for advanced economies shows univocally that surveys, which provide the most timely information, contribute to an improvement in the nowcasting in the early part of the quarter, before hard data like industrial production and retail sales become available (Matheson *et al.*, 2007; Angelini *et al.*, 2008; Darne, 2008; Frale *et al.*, 2010). However, once the latter are released the contribution of these survey vanishes. In contrast to these findings, our results show mixed evidences on the ability of survey data available for India to enhance the predictive accuracy of our nowcasts. To date, among survey data only the Purchasing Managers Index series are available with a monthly frequency (both for manufacturing and the services sector), lbeit only from 2007. The other important source of survey data is the business survey conducted by the Reserve Bank of India. We find that it enhances only marginally the predictive accuracy of of our targets growth rate with respect to the benchmark models, but hardly improves the models containing more timely real indicators such as IIP manufacturing.[1] We relate this finding to the fact that the RBI business survey is released to the public only quarterly, and with a small time advantage with respect to the GDP release. This significantly reduces its usefulness to nowcast GDP, as hard data are already available covering most of the reference quarter.

We should also stress that, ideally, our exercise, to be truly *real-time*, should properly take into account the entire history of data releases of the national accounts series (to a lesser extent of the monthly proxies). India's quarterly GDP are subject to periodic revisions along with the annual estimates of GDP, that embody more accurate information regarding the economy (e.g. estimates of the informal sector). These revisions influence the nowcasting performance of our current exercise. This occurs because the quarterly GDP series which we use as a *target*, especially in the earlier part of the sample, refer to a revised GDP figure. Access to real time vintages of the individual data releases (not available to date) would probably reinforce the robustness of our results.

The paper is organised as follows. *Section 2* describes the target of our nowcasting exercise. *Section 3* outlines various models underlying the forecast evaluation mechanism and the framework for the forecasting exercise in the pseudo real-time. *Section 4* describes the data used for the analysis highlighting various survey data available in India and the nature of information flow in real-time in the economy. *Section 5* presents results of our pseudo-real time exercise of nowcasting GDP growth and *Section 6* concludes.

[1]This result applies not only for the survey responses regarding the current quarter assessment, but also to those referring to expected movements in the following quarter.

Figure 1: GDPXagri, GDPXoth and other service value added Growth



This figure show year-on-year growth rate of components of GDP and other services value added.

2. What we are Tracking: GDP Growth

The estimation of Gross Domestic Product is the result of a complex statistical procedure drawing on multiple data sources. It generally relies on rigorous models as well as more *ad-hoc* routines. Most countries compile national accounts at an annual and a quarterly frequency. At the quarterly frequency the procedure is simpler, as the information available to the statistician is limited. Nevertheless, in the latter case the challenge for the statistical offices is to infer from the available sources, a timely picture of the economy and to properly embed this within the more exhaustive information that becomes available when the annual accounts are compiled.

We attempt to observe the business cycle movements of the economy. We choose the sub-components of GDP as the target which are directly related to the business cycle fluctuations in the economy. A large part of GDP is still driven by the fluctuations in agricultural output, clustered in two quarters of the year when the main crops are harvested. Despite the declining weight of agriculture in overall GDP, bad-crop years can lead to marked swings in the year-on-year growth rate of overall GDP (an example is 2002). Timely information on the developments in agricultural output and reliable crop estimates are not as easily available as other economic data. As the factors underlying agricultural output (rainfall, temperature, etc) are probably different from the ones driving fluctuations in the rest of the economy we decided to choose GDP excluding agriculture as one of our target variables. The left panel of *Figure 1* shows year-on-year (henceforth YOY) growth rate of GDP *vis-a-vis* GDP excluding agriculture. It shows larger fluctuations in the former compared to the latter. Moreover, the growth rate of GDP seems to be orthogonal to the growth rate of GDP excluding agriculture at various intervals during Q2-2000 to Q2, 2010.

In addition, the sector “other services”, mainly composed of government services is subject to significant short run volatility due to the dynamics of public sector outlays. For instance, the right panel in *Figure 1* shows a huge jump in the growth rate of the other services in Q4-2008. This is precisely due to the increase in public sector wages following the implementation of the Sixth Pay Commission Report. This large movement in the short run may add to the volatility in the growth rate of the overall GDP in the short run. Hence we choose GDP excluding both agriculture and other services as another target variable for our analysis (*see the right panel of Figure 1)*

3. Models

This section describes different models used for forecasting GDP growth. We consider models that utilise only quarterly data as well as the models that exploit information from monthly data. The models are designed to be used in real time and that at each date of the forecast some of the proxy series, due to publication lags, will have missing data at the end of the sample. Moreover, due to the different timing of data releases, the number of missing data differs across series. Missing data will be forecasted using simple univariate monthly autoregressive models.

3.1 Quarterly Models

3.1.1 Naive random walk and auto-regressive models

We use two univariate time series models for quarterly YOY growth rate of GDP (g_t^Q) as benchmark models. These are:

a) Naive model (random walk forecast)

$$g_t^Q = g_{t-1}^Q + \varepsilon_t^Q \quad (1)$$

b) First-order autoregressive (AR) model

$$g_t^Q = ?g_{t-1}^Q + e_t^Q \quad (2)$$

where ε_t^Q is quarterly white noise with zero mean and a constant variance σ_ε^2.

In addition to using absolute measure of forecast performance, we evaluate the forecast performance of different models relative to these benchmark models.

3.2. Bridging Monthly Data or Survey Variables with Quarterly GDP

Bridge equations are a widely used method to forecast quarterly GDP growth using information from various other indicators. The set of indicators comprises of real activity variables, financial variables, survey variables and international survey and activity variables.

3.2.1 Auto regressive models with exogenous proxies (AR+X models)

Let us denote growth (YOY) in our quarterly target variable as g_t^Q and the vector of k selected monthly indicators, for every AR+X model j, as $x_t^j = (x_{1,t}^j, \ldots\ldots, x_{1,t}^j)'$,

t=1,……,T. The models are estimated from quarterly aggregates of the monthly data. Predictions of the target GDP growth series are obtained in two steps. In the first step, the monthly indicators are forecasted over the remainder of the quarter to obtain forecasts of their quarterly aggregates, $x_{k,t}^{j,Q}$. The forecasts of the monthly predictors are based on univariate time series models, using an automatic model selection relying on the AIC information criterion. In a second step, growth rates of the resulting quarterly aggregates are used as regressors in the bridge equation to obtain the GDP forecast, with the following structure:

$$g_t^Q = \mu + \phi_2 g_{t-1}^Q + \sum_{i=1}^{k} \beta_i^j(L) g_{xi,t}^{j,Q} + \varepsilon_t^{j,Q} \qquad (3)$$

where μ is an intercept parameter and $\beta_i^j(L)$ denotes a lag polynomial. The variable $g_{xi,t}^{j,Q}$ is the YOY growth rate of quarterly aggregate of monthly indicator x_i used in model j.

3.2.2 Bridge bottom up

The forecast of the growth in our quarterly target variable, g_t^Q, is obtained *indirectly* by aggregating the sectoral growth rates constituting the target component of GDP. The latter are in turn obtained using specific monthly indicators which act as a proxy for the development in a given sector, Let $\{x_{1t}, x_{2t}, \ldots x_{kt}\}$ be this set of monthly indicators. Therefore we start from a set of R sectoral value added AR+X equations, just like in *Equation 3.2.1*, where monthly proxies are first forecasted to reach the end of the quarter and then aggregated to the quarterly frequency.

$$VA_r^Q, t = \mu + \phi_2 VA_{r,t-1}^Q + \sum_{i=1}^{k} \beta_i^r(L) g_{xi,t}^{r,Q} + \varepsilon^{r,Q} \qquad (4)$$

where r=1,…, R are the R sectors making up the target component of GDP.

Finally, the growth rate of target variable, g_t^Q, is obtained by aggregating the predicted sectoral value added growth rates VA^Q using the sectoral weights in the target component of GDP. The sectoral weights at period *t-4* are used to obtain aggregate growth prediction at period t:

$$g_t^Q = \sum_{r=1}^{R} \gamma_{rt-4} \hat{VA}_{rt}^Q \qquad (5)$$

where γ_r denotes the share of rth sector in the target component of GDP. This approach of obtaining overall growth rate by aggregating the monthly proxies for various sub sectors of GDP is referred to as the *bottom up approach*.

3.2.3 Bridging with factors

The quarterly GDP growth can be bridged with the factors obtained from monthly information:

$$g_t^Q = \mu + \beta' f_t^Q + \varepsilon_t^Q \tag{6}$$

where f_t^Q is the common factor driving quarterly aggregates of all the monthly indicators.[2] Given a set of monthly time series, $x_t=(x_{it},........,x_{nt})$, the factor structure is given by:

$$x_t = \Psi f_t + \eta_t \tag{7}$$

This equation relates the $n\times1$ vector of monthly time series x_t to the $p\times1$ vector of common factors $f_t=(f_{1t},......f_{nt})'$ through a matrix of loadings Ψ and to the idiosyncratic component $\eta_t=(\eta_{1t},.......,\eta_{nt})$. The number of static factors p is smaller than the number of series n. The factors are extracted from the growth rate of quarterly aggregate of monthly indicators with the application of principal component analysis. Finally, the forecast of GDP growth is obtained using the bridge *Equation 6*.

3.2.4 The Pseudo Real-time Forecasting Exercise

We design the forecast evaluation exercise to predict quarterly growth of two alternative targets: GDP excluding agriculture and GDP excluding agriculture and other services using monthly indicators which are published within the quarter.

We conduct our out of sample forecast exercise over a period of six years from Q2-2005 to Q2-2010. In the process, we attempt to replicate the real-time application of the models by mimicking the real-time pattern of the data release. The parameters of the models are estimated in a recursive manner exploiting the information available only at the time of forecast. We do not have a real-time database for all the predictors considered, therefore we will not be able to take into account the real-time data revisions. Instead, we use a data set downloaded on August 31, 2010 and combine this with the typical data release calendar to reconstruct data availability at the end of each month.

For GDP of a given quarter, we produce a sequence of forecasts in three consecutive months prior to the release of the official quarterly GDP. We will label these three sequences, as *month-2*, *month-1*, *month-0*, respectively denoting the forecasts two months from the GDP release, one month, and a few days ahead from of it. Starting from the N variable dataset extracted on August 31, 2010 (T), $O_t=\{x_s\}^T_{s=1,}$ we define a pseudo real-time dataset $\Omega_t=\{x_s\}_{s=1}^t$ as the observations from the original dataset $\Omega_T=\{x_s\}_{s=1}^T$, but with observations $x_{j,t-h}$, $h\geq0$ and $j=1,....$, N if observations $x_{j,T-h}$ are missing in $\Omega_T=\{x_s\}_{s=1}^T$.

---
[2]We use a dataset of approximately 50 monthly time series.

In order to deal with the missing observation of monthly indicators at the end of the sample, we forecast $x_{j,t-h}$ h step ahead to fill information till the end of the quarter. Then x_{jt}, t=1,2,3 ? t_Q is aggregated over the quarter to obtain quarterly value of the monthly indicator x_{jt}^Q where t_Q indicates a particular quarter. A forecast $g^Q_{t_Q|l}, l > t_Q$ made in l,l=2,1,0 months ahead of the release of GDP number for the quarter t_Q is based on information set Ω_{t_m} where t_m, m=1,2,3 denotes the set of months spanning the quarter t_Q.

More precisely, the parameters under different models are estimated recursively using existing information till g^Q_{t-1Q} and O_{t-1m}. Using the estimated parameters, the forecast of $g^Q_{t_Q}$ is then obtained by exploiting the information available for Ω_{t_m}.

4. Data

4.1 The Indian Context

In India the Central Statistical Organization (CSO) introduced the quarterly estimates of GDP in 1999, both at current and constant prices as part of the requirements under the Special Data Dissemination Standard of the IMF.[3] Currently, the quarterly figures, dating back to 1996, become available in Datastream, the multi-country database maintained by Thomson Reuters. In this paper, we use the time series of GDP from Datastream. In India, the quarterly GDP releases with a delay of approximately two months with respect to the end of the reference period: for instance, the data for Q4-2009[4] were published on February 26, 2010.[5]

The CSO also produces the breakdown into sectoral value added and into the main demand side components. The supply side estimates, i.e. those obtained by summing the value added of the different kind of activities, are deemed to be more reliable because of the large set of underlying indicators used in the estimation. The time series of sectoral components of GDP are also taken from *Datastream* for our analysis.

4.2.1 Production Side Estimation from Monthly Variables

The quarterly estimates from the production side are based on the so called *benchmark-indicator* approach. In particular, for each of the industry groups, a set of (mainly) physical indicators on which data is available at quarterly (or higher frequency) is used to extrapolate the value added in the reference sector from the same quarter of the previous year.

A fairly detailed description of the main indicators employed as proxies by the CSO is well documented in the national accounts manuals (CSO, 2007), however a

[3] *See*, http://dsbb.imf.org/pages/SDDS/DQAFBase.aspx?ctycode=IND&catcode=NAG00
[4] Throughout the paper, Q1 refers to January-March, Q2 refers to April-June, Q3 refers to July-September, and Q4 refers to October-December.
[5] See http://mospi.nic.in/mospi_press_releases.htm for the most recent releases.

certain margin of uncertainty remains in the exact methods and in the way the indicators are used to estimate quarterly GDP.

Indeed, the official estimation of GDP always remains to a certain degree not replicable, even ex-post. First, because not all the information set available to the statistical office is made public, for confidentiality reasons or simply because of the information advantage that the CSO has over its own statistics. Second, because some details in the procedures used by the CSO will not be entirely replicable.

In this section we attempt to reconstruct quarterly GDP[6] growth, from a small set of monthly indicators, for each sectoral value added at different time points prior to its release.

Table 1: Indicators used for Quarterly Estimates of GDP Growth

This table shows monthly indicators used for reconstructing quarterly estimates of GDP growth. The set of indicators is a subset of indicators used by CSO to estimate quarterly Indian GDP.

Sectors	**Indicators**
Mining and quarrying	IIP mining, monthly production of coal and crude petroleum
Manufacturing	IIP manufacturing
Electricity, gas and water supply	IIP electricity
Construction	Monthly production of cement, steel and coal
Trade, hotels, transport and communication	Commercial vehicles production, railway goods traffic, port traffic, cellular subscription
Banking and insurance	Deposits, non food bank credits, WPI, NSE turnover
Other services	Central govt revenue expenditure net of interest payments, CPI

Table 1 shows the monthly indicators used for reconstructing quarterly estimates of GDP growth. While we do not have access to some of the indicators used by the CSO, we do consider some monthly indicators that we think might have some impact on the sectoral value added. As an example, we consider turnover on the NSE as one of proxy indicators for GDP (banking and insurance).[7] The methodology essentially relies on bridge equations, developed to link early monthly releases with quarterly GDP growth for each sectoral value added. The monthly indicators except for steel data are sourced from the Business Beacon database produced by the Centre for Monitoring Indian Economy (CMIE). The data on steel production is sourced from *Datastream*.

[6]Henceforth, by GDP we mean either GDPXagri or GDOXoth.
[7]Although we find that this indicator does not provide significant information for improving the model fit.

4.2.2 Information from Surveys

In addition to the monthly variables used by the Statistical Office, the survey variables can also provide valuable information about the state of the economy. Using survey data to nowcast GDP growth has some inherent advantages. These are: (i) survey data provide a signal that is obtained directly from the participants regarding the short-term evaluation of their activity; (ii) they are more timely than the hard data; and, (iii) they are subject to less revisions. But unlike hard data, they are based on sentiments and expectations and are sensitive to sample size and composition. A number of papers (Angelini *et al.*, 2008; Matheson *et al.*, 2007), investigate the forecast performance of survey data to nowcast GDP. Giannone *et al.* (2009) find that due to their timely nature, surveys provide valuable information and the early signal that they provide can be considered as a reliable indicator of economic conditions before hard indicators are released. Matheson *et al.* (2007) find that exploiting the panel dimension to qualitative survey data can give a better signal about official data.

In India the usefulness of business survey data has never been evaluated in an out of sample exercise. The Reserve Bank of India routinely describes their trends in the *Outlook* chapter of the *Macroeconomic and Monetary Developments* quarterly publication. We focus on three types of surveys:

- RBI business survey
- The Market Purchasing Managers Index (PMI) for India as well as the JP Morgan World Business Survey
- The Dum and Bradbury (D&B) composite business survey

The Reserve Bank of India has been conducting Industrial Outlook Surveys, since 1998 on a quarterly basis with a view to gain insight into the performance and prospects of the private corporate sector engaged in manufacturing activities. The survey is released at the end of each quarter with the RBI's publication on *Macroeconomic and Monetary Developments.* As an example, the results of the 50th round of the Industrial Outlook Survey for April-June, 2010 was released on July 26, 2010. It provides an assessment for April-June quarter and expectations about the next quarter (July-September) for a host of variables affecting the industrial and economic environment.[8]

The Purchasing Managers Index(PMI) is released on a monthly basis. Its global index is released by JP Morgan, while the Indian survey is conducted by HSBC and Markit Economics, both for the manufacturing and the services sector. The HSBC PMI manufacturing index is based on a survey of 500 companies. The index compiles a variety of factors such as output and employment growth, pricing pressures, order flow and delivery lags, among other indicators. A reading of over 50 indicates expansion in this indicator. The PMI survey data are released at the end of the month. For instance the release date of this indicator for the month of July is August 2, 2010. The PMI survey

[8]Specifically, the variables are: overall business situation, financial situation, working capital finance requirement, availability of finance, production, order books, cost of raw materials, inventory of raw materials, inventory of finished products, capacity utilisation, level of capacity utilisation (compared to the average in four quarters), assessment of the production capacity with regard to expected demand in the next six months, employment in the company, exports, imports, selling prices, increase in selling prices and profit margin.

data for India are available only from 2007. Hence, for the purpose of our analysis we backcast the series using IIP manufacturing and business cycle indicator for EU.

The *D&B* Business Optimism Index for India is well known among investors and policy-makers. The survey is released a few days after the end of each quarter. The index is formed on the basis of a quarterly survey of business expectations. It is conducted on a sample of companies that are selected randomly from the *D&B* commercial credit file, and includes both the manufacturing and the services sectors. A composite Business Optimism is obtained as a weighted average of six questions on business developments over the past and next year.[9]

4.3 The Calender of Real-time Data Flow

Table 2 shows a snapshot of our information set $\Omega_T = \{x_s\}_{s=1}^{T}$, together with the most recent release dates for every variable. It shows the typical "jagged edge" shape determined by the non-synchronous nature of the Indian data releases.

Table 2: Data available on August 31, 2010, just ahead of Q2-2010 GDP release

This table presents a snap-shot of the information flow in real time in the Indian economy.

	March -10	**April – 10**	**May – 10**	**June - 10**	**July - 10**	**August-10**	**Last release**
IIP	x	x	x	x	x	x	12 Aug 2010
Cement production	x	x	x	x	x	x	26 Aug 2010
Steel production	x	x	x	x	x	x	26 Aug 2010
Coal production	x	x	x	x	x	x	26 Aug 2010
Railway goods traffic	x	x	x	x	x	x	26 Aug 2010
Port traffic	x	x	x	x	x	x	26 Aug 2010
Tourists arrivals	x	x	x	x	x	x	27 Aug 2010
Vehicles production	x	x	x	x	x	x	11 Aug 2010
Electricity	x	x	x	x	x	x	02 Aug 2010
Cellular subscription	x	x	x	x	x	x	13 Aug 2010
Nonfood bank credits	x	x	x	x	x	x	14 Aug 2010
Deposits	x	x	x	x	x	x	14 Aug 2010
Government expenditure	x	x	x	x	x	x	28Aug 2010
BSE	x	x	x	x	x	x	31 Aug 2010
PMI surveys	x	x	x	x	x	x	6 Aug 2010
US IIP	x	x	x	x	x	x	5 Aug 2010

On August 31, 2010, with a delay of two months with respect to end of the reference quarter (March 31, 2010), the second quarter GDP (calendar year) was released. At that date the index of industrial production (IIP) was available up to June, 2010, having been released on August 12. On the other hand, the information flow on commercial vehicles production is more timely: on August 31, 2010 data up to July was available, having been released on August 11, 2010.

[9]The questions regard net sales, net profits, selling prices, new orders, inventories, and employee levels.

5. Results

In this section, we discuss the key findings from our pseudo real time tracking exercise and evaluate the forecast performance of the various models over the period from Q2-2005 to Q2-2010.

5.1 Forecast Evaluation

This table presents RMSE from various forecast models for the target GDP excluding agriculture.

Table 3: RMSE from forecast models for GDPXagri

Model	**Forecast**	**Nowcast**		
		Month 0	**Month 1**	**Month 2**
Naive	1.054			
AR	1.066			
Bridge bottom up		**0.816**	**0.825**	**0.898**
Manufacturing		1.117	1.485	2.185
Electricity		0.948	1.336	1.773
Construction		2.507	2.507	2.57
Mining		2.321	2.498	2.313
Service		1.591	1.591	1.595
Finance		1.414	1.414	1.465
Other Services		3.047	3.955	5.508
AR+X Models				
IIP manufacturing		1.21	1.144	1.14
PMI global		1.062	1.062	1.052
US IIP		1.114	1.114	1.142
PMI India		1.118	1.118	1.126
Bradbury survey	0.973			
RBI survey business situation	0.978			
RBI survey financial situation	0.967			
Factor model		1.163	1.136	1.139

We focus separately on the results relating to the two target variables. As to our first target, non-agriculture GDP (GDPXagri), the results are reported in *Table 3.*

What follows next, is an attempt to summarise how flow of information over time and across sectors helps us to gauge the growth rate in GDPXagri. Starting from the naive and AR models, which by definition contain no information regarding the current quarter to be nowcasted, we find that the RMSE of the forecasts is of approximately 1.06 percentage points. The performance of the model does not improve even if we start adding information (moving from month-2 to month-0) using a single indicator (AR+X models) where the set of indicators consists of IIP manufacturing, PMI for India as well as the JP Morgan World Business Survey and US IIP. Interestingly, the quarterly RBI survey data, while improving the RMSE by 8.25 percent on average, compared to the benchmark models, seem to contain additional predictive content with respect to the models containing more timely real indicators such as IIP manufacturing. Also, the quarterly Dum and Bradbury (D&B) composite business survey seems to improve the predictive power of the benchmark AR model.

Finally, we find that, for the three consecutive months considered in our exercise, the bottom-up bridge model outperforms not only the benchmark models, but also all the single indicator models. The RMSE drops by 8.89 percent from 0.90 at month 2, to 0.82 at month-0. We also find that out of sample growth rate in some sectors is harder to predict using the indicators available. This can be gauged by comparing the RMSE of the individual sub-sectors when new information is added to the models. In particular, more complete information on the developments in the proxy variables for the mining sector (IIP mining and coal production) does not lead to a reduction of the RMSE of our forecasts. Moreover, it appears that the first stage model for the "other services" component performs very poorly in terms of forecast accuracy; this, in turn, reduces the overall forecast accuracy of our second stage bottom-up model. This result suggests that our procedure is departing somewhat from the one adopted by CSO to estimate the growth of value added in the public sector. In particular, our information set for this sector may not be exhaustive with respect to the CSO information set. This intuition is confirmed by investigating the history of the individual forecast errors. These are particularly large in the last part of the sample when government outlays, in connection with the Sixth Pay Commission and the stimulus packages, recorded a sudden abnormal behaviour (*see* the right panel of *Figure 1.)*

This however does not impinge on the overall ability of the bottom-up approach to provide, not only a more accurate forecast for our target, but also a more informative view of the contributions of each sector to a given forecast. Most interestingly, bridge model outperforms the factor model in terms of forecast accuracy. This result is consistent with the fact that the narrow information set used by the CSO for the construction of GDP is orthogonal to the large number of real time variables that capture several components of the activities in the economy. This result is consistent with Matheson, 2011. The author nowcasts GDP of a large number of countries using various models. He finds that for India, along with Australia and Saudi Arabia, factor model based on a large data set performs poorly compared to simpler models such as pooled bridge model.

Table 4: RMSE from forecast models for GDPXoth

Model	Forecast	Nowcast		
		Month 0	**Month 1**	**Month 2**
Naïve	1.302			
AR	1.304			
Bridge bottom up		**0.728**	**0.733**	**0.858**
Manufacturing		1.117	1.485	2.185
Electricity		0.948	1.336	1.773
Construction		2.507	2.507	2.57
Mining		2.321	2.498	2.313
Service		1.591	1.591	1.595
Finance		1.414	1.414	1.465
Other Services		3.047	3.955	5.508
AR+X Models				
IIP manufacturing		0.887	0.898	1.22
PMI global		1.219	1.219	1.22
US IIP		1.451	1.451	1.471
PMI India		1.39	1.39	1.423
Bradbury survey	1.111			
RBI survey business situation	1.303			
RBI survey financial situation	1.082			
Factor model		1.156	1.206	1.237

This table presents RMSE from various forecast models for the target: GDP excluding agriculture and other services.

Table 4 presents RMSE of alternative models when the target variable is GDP excluding agriculture and other services (GDPXoth). For this alternative target too, bridge bottom up model outperforms the benchmark models as well as the AR model augmented with monthly and quarterly indicators. However, for this target, the simpler bridge model relying only on IIP manufacturing also performs satisfactorily. Again, unlike the GDPXagri, the Indian survey data, neither improve the RMSE compared to the benchmark models, nor seem to contain additional predictive content with respect to the models containing more timely real indicators.

We report the RMSE of each model relative to the benchmark AR model in *Table 5*. It shows the average RMSE of forecasts at two, one and zero month ahead of GDP release relative to the forecast from the benchmark AR model when the underlying models utilise information flow from monthly indicators. A number lower than one indicates that the model's forecasts are more accurate than the average growth over the past sample. The findings differ qualitatively among the two alternative targets. For the target GDPXagri, bridge bottom up followed by the AR+X models with PMI global indicator and all quarterly survey variables outperform the benchmark AR model. For the alternative target GDPXoth, except for AR+X models with PMI India survey, US IIP and RBI survey on business situation, all other alternative models out perform the benchmark AR model.

Table 5: Average RMSE Relative to the Benchmark

This table shows average RMSE of forecast models relative to the benchmark AR model.

Model	Average RMSE	
	GDPXagri	GDPXoth
Naïve	0.989	0.999
Bridge bottom up	0.794	0.593
AR+X models		
IIP manufacturing	1.114	0.683
PMI global	0.993	0.935
US IIP	1.045	1.113
PMI India	1.049	1.066
Bradbury survey	0.913	0.853
RBI survey business situation	0.917	1
RBI survey financial situation	0.907	0.83
Factor model	1.082	0.9

5.2 Forecast Model Encompassing Tests

The encompassing test between bridge and other alternative models is based on a regression of the actual data y_t^Q on forecasts $f_{bridge,t}^Q$ and $f_{alternative,t}^Q$ from two models,

$$y_t^Q = \lambda f_{bridge,t}^Q + (1-\lambda) f_{alternative,t}^Q + u_t, \quad 0<\lambda<1$$

We test the null that $\lambda=1$ implying that bridge dominates other models, using the Newey-West corrected standard error. The acceptance of the null hypothesis indicates that alternative models can not provide additional information over and above the information provided by the bridge model to improve forecast performance. The test results are reported in Table 6. We find that when GDPXagri is the target, the null hypothesis is rejected for all alternative models except for the factor model and the AR+X models using PMI global survey and IIP manufacturing as indicators. This implies that except for these models, all other alternative ones are able to provide additional information improving forecast performance. However for the target variable GDPXoth, the null hypothesis is accepted for all alternative models indicating that bridge bottom up dominates these models.

This table presents test results on the predictive power of alternative models *vis-a-vis* bottom up bridge as the reference model. The null hypothesis of the test is that the bottom up bridge contains the entire predictive power, i.e., $\lambda=1$.

Table 6: Alternative Models *vis-a-vis* Bridge as the Reference Model

Model	GDPXagri		GDPXoth	
	Estimate	**p-value**	**Estimate**	**p-value**
Naïve	0.658	0.017	0.869	0.404
AR	0.664	0.026	0.875	0.431
AR+X models				
IIP manufacturing	0.911	0.615	0.924	0.748
PMI global	0.642	0.061	0.926	0.672
US IIP	0.657	0.024	0.921	0.576
PMI India	0.674	0.027	0.906	0.531
Bradbury	0.609	0.015	0.758	0.112
RBI survey business situation	0.588	0.023	0.906	0.572
RBI survey financial situation	0.548	0.025	0.819	0.38
Factor model	0.775	0.257	0.957	0.743

Alternatively, we test the null that $\lambda=0$, using Newey-West corrected standard error. The rejection of the null hypothesis will imply that information provided by the bridge bottom up model helps in improving the forecast performance. The test results are reported in *Table 7*. We find that for both the targets, null hypothesis is always rejected, indicating that the information from bridge bottom up always improve forecast performance.

This table presents test results on the predictive power of alternative models *vis-a-vis* bottom up bridge as the reference model. The null hypothesis of the test is that the bottom up bridge contains no predictive power, i.e., $\lambda=0$.

Table 7: Alternative Models *vis-a-vis* Bridge as the Reference Model

Model	GDPXagri		GDPXoth	
	Estimate	p-value	Estimate	p-value
Naive	0.658	0	0.869	0
AR	0.664	0	0.875	0
AR+X models				
IIP manufacturing	0.911	0	0.924	0.015
PMI global	0.642	0	0.926	0
US IIP	0.657	0	0.921	0
PMI India	0.674	0	0.906	0
Bradbury	0.609	0	0.758	0
RBI survey business situation	0.588	0.001	0.906	0
RBI survey financial situation	0.548	0.001	0.819	0
Factor model	0.775	0	0.957	0

6. Conclusions

This paper applies bridge and factor models to nowcast short-term GDP growth in India. The methodology is designed to “bridge” early releases of monthly indicators to quarterly GDP. A bottom up approach is followed where for each sub sector of GDP, relevant monthly indicators are identified and bridge models are estimated on the year-on-year growth rate of quarterly value of monthly variables to predict year-on-year growth rate of GDP.

The bridge models are applied in a pseudo real-time setting- by actually taking into account the information set available at each point in time to nowcast GDP growth. The nowcasting exercise is conducted at three intervals: two months, one month and few days before the actual GDP release. The results of the nowcasting exercise show that bridge bottom up model significantly outperform alternative models including the factor model.

Finally, we investigate for the first time the effective usefulness of survey data available for India in nowcasting GDP. Our results suggest that Indian survey data enhance only marginally the predictive accuracy of our nowcasts. As to the Reserve Bank of India business survey, given its quarterly nature, and its small time advantage with respect to the GDP release, we find that it is of little use to improve our nowcasts of GDP since hard data are already available covering most of the reference quarter. As to the PMI survey, we find that, despite its timeliness, it does not improve the nowcasting of the benchmark AR and Naive models. These findings are in stark contrast with those found in developed economies, where survey dynamics are largely consistent with the one recorded by the official GDP growth rate. This peculiarity may arise from a marked difference in coverage in the reference sample of firms underlying the GDP figures calculated by the CSO (especially for industry and private services) and the ones considered both by the RBI and the PMI surveys.

References

Angelini E, *et. al.,* 2008. "Short-term forecasts of Euro Area GDP growth." European Central Bank Working Paper No. 949.

Baffigi A, R. Golinelli, G. Parigi, 2003. "Bridge models to forecast the euro area GDP." *International Journal of Forecasting*, 20(3); 447–460.

Banbura M, Giannone D, 2010. "Nowcasting" ECARES working paper, No. 2010-021.

Barhoumi K, *et. al.,* 2008. "Short Term Forecasting of GDP using large datasets: a pseudo-realtime forecast evaluation." Occasional Paper Series, European Central Bank.

Borin A, R. Cristadoro, R. Golinelli, G. Parigi, 2010. "Forecasting World Output: the Rising Importance of Emerging Asia." http://www.bancaditalia.it/studiricerche/convegni/atti/chines% e-economy/sessione1/borini/Borin_1.pdf .

CSO, 2007. "National Account Statistics: Sources and Methods." *Technical report*, Central Statistical Organisation, Ministry of Statistics and Programme Implementation.

D'Agostino A, K, McQuinn, D. O'Brien, 2008. "Now-casting Irish GDP." *Research Technical Paper, No. 9/RT/08*, Central Bank and Financial Services Authority of Ireland.

Darne O., 2008. "Using Business Survey in Industrial and Services Sector to Nowcast GDP Growth:The French case." *Economic Bulletin*, 3(32).

Frale C, M Marcellino, G. Mazzi, T. Proietti, 2010. "Survey Data as Coincident or Leading Indicators." *Journal of Forecasting*, 29.

Giannone D, L. Reichlin, S. Simonelli, 2009. "Nowcasting Euro Area Economic Activity in Real-Time: The Role of Confidence Indicators." ECARES Working Paper 2009-021.

Giannone D, L. Reichlin, D. Small, 2008. "Nowcasting: The Real-Time Informational Content of Macroeconomic Data." *Journal of Monetary Economics*.

Knut AA, T. Trovik, 2007. "Nowcasting Norwegian GDP: The role of asset prices in a small open economy." Working Paper, No. ANO/2007/9, Norges Bank.

Maier P., 2011. "Mixed Frequency Forecasts for Chineese GDP." Working Paper, No. 2011-11, Bank of Canada.

Matheson T., 2011. "New Indicators for Tracking Growth in Real Time." *Working Paper, No. WP/11/43*, International Monetary Fund.

Matheson T, J. Mitchell, B. Silverstone, 2007. "Nowcasting and Predicting Data Revisions in Real Time Using Qualitative Survey Data." Reserve Bank of New Zealand, Discussion Paper 2007/02.

Pedersen M., 2010. "Extracting GDP Signals from the Monthly Indicator of Economic Activity: Evidence from Chilean Real-Time Data." Working Paper, No. 595, Central Bank of Chile.

Winters L, S. Yusuf, 2007. *Dancing with Giants*. The World Bank and The Institute of Policy Studies.
http://books.google.co.in/books?hl=en&lr=&id=foqqed2TPz8C&oi=%fnd&pg=PR7&dq=India%27s+integration+with+world+economy&ots=YR1DbmUHl8&sig=Q2ku%nyOIrxPEPtrmlMOsr0QiE3s#v=onepage&q=India%27s%20integration%20with%20world%20e%conomy&f=false.