Tracking Indian growth in real time *

Rudrani Bhattacharya, Radhika Pandey and Giovanni Veronese[1]

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 policy maker at each point in 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. Unlike in advanced economies, Indian survey data are found to provide little additional information with respect to the hard data already used in the national accounts.

[1]Bank of Italy and NIPFP. Correspondence: giovanni.veronese@esteri.it

*We gratefully acknowledge encouragement and support from Ajay Shah, Ila Patnaik and the team at National Institute of Public Finance and Policy. The views expressed here do not necessarily reflect those of the Bank of Italy

1 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., 2004) the modelling has become increasingly complex to account for the larger information sets available nowadays, and to properly formalize 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 (see Giannone, Reichlin and Small 2008).

Tracking growth in the Indian economy would be best performed using a measure like GDP. Unfortunately official estimates of this indicator are released with considerable delay, suffer from sizeable revisions and are not available in seasonally adjusted format. The first official release of quarterly GDP growth is published approximately 7-8 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 idyosincratic noise.

Some of these indicators are themselves part of the inputs into the quarterly national account computations performed by the Central Statistical Organization (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.

Some of these sectors' developments are unrelated to the business cycle movements of the economy, and display considerable volatility. In particular, agriculture is 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. For this reason, in the following we restricted our nowcasting analysis to two measures of growth, based on GDP-excluding agriculture (GDPxagri) and GDP excluding agriculture and other services (GDPxoth).

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.

Our findings show that an effective nowcast of GDP in the Indian context can be performed by using simple bridge models, that either rely on single indicators (e.g. industrial production, global survey data, etc.) or more elaborate models that strive to mimick as closely as possible the national accounts estimation procedure.

We find that bridge models [Baffigi et al., 2003, Barhoumi et al., 2008], that are regressions of quarterly GDP growth on a small set of preselected key monthly indicators, serving as proxies for the various sub-sectors of the economy perform satisfactorily in predicting current quarter 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. However, the performance of the bridge methodology is not as satisfactory when predicting the GDPxagri growth. The inherent difficulty to predict developments in the "other services" series, negatively affects the forecasting power of our disaggregate bridge model, while the one from simpler models remains unaltered.

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.

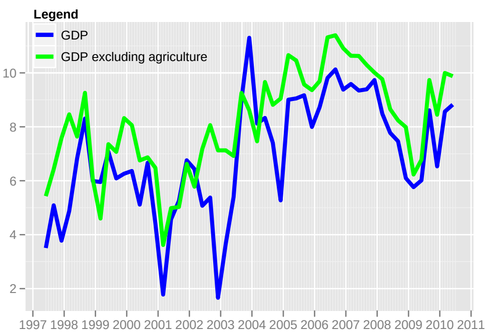
Finally, we investigate for the first time the effective usefulness of Indian survey data in nowcasting GDP. The literature for advanced economies shows univocally that surveys, which provide the most timely information, contribute to an improvement in the nowcasting in the early part of the quarter, before hard data like industrial production and retail sales become available (Angelini, Camba-Mendez, Giannone, Runstler and Reichlin. 2008). However, once the latter are released the contribution of these survey vanishes. In contrast to these findings, our results suggest that the survey data available for India is not enhancing 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), albeit only from 2007. The Reserve Bank of India business survey is instead released quarterly, and with only 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. This finding applies not only for the survey responses regarding the current quarter assessment, but also for those referring to expected movements in the following 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 data release would probably reinforce the readiblity of our results.

The paper is organized as follows. Section 2 describes the target of our nowcasting exercise, and sheds some light on the current methodology to construct the GDP estimates by the CSO. It then describes the survey data available in India. Section 3 presents our pseudo-real time exercise of nowcasting GDP growth. Sections 4 concludes.

Figure 1 YOY growth rate: GDP and GDPxagri

This figure shows YOY growth rate of GDP and non-agricultural GDP.



2 What we are tracking: GDP growth

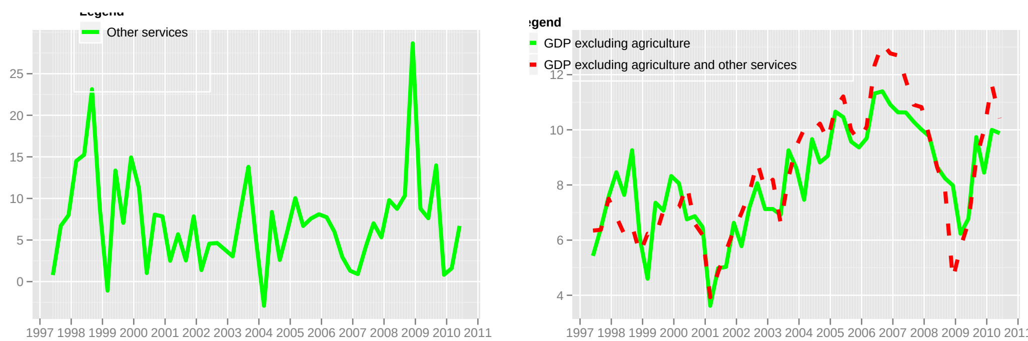
The estimation of Gross Domestic Product is the result of a complex statistical procedure drawing on multiple data sources. It generally relies on rigorous models as well as more ad-hoc routines. Most countries compile national accounts at an annual and a quarterly frequency. At the quarterly frequency 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 it 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 crops 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 rather 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 (see Figure 1).

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 left panel in Figure 2 shows a huge jump in the growth rate of the other services in 2008 Q4. This is precisely due to the implementation of the sixth pay commission reports. This large moments in the short run may add to the volatility in the growth rate of the over all GDP in the short run. Hence we choose GDP excluding both agriculture and other services as another target variable for our analysis (see right panel of Figure 2).

Figure 2 YOY growth rate: other services GDPxoth

This figure shows YOY growth rate of other services and GDPxoth vis a vis GDPxagri



2.1 The Indian context

In India the Central Statistical Organization (CSO) introduced in 1999, the quarterly estimates of GDP, both at current and constant prices as part of the requirements under the Special Data Dissemination Standard of the IMF.[1]. Currently, the quarterly figures, dating back to 1996, become available with a delay of approximately 2 months with respect to the end of the reference period: for instance, the data for Q4-2009[2] were published on 26th February, 2010 (see here for the most recent releases).

[1]See, `http://dsbb.imf.org/pages/SDDS/DQAFBase.aspx?ctycode=IND&catcode=NAG00`

[2]Throughout the paper Q1 refers to Jan-March, Q2 refers to Apr-June, Q3 refers to Jul-Sept and Q4 refers to Oct-Dec.

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 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. This process can be sketchedly formalized as:

$$\begin{aligned} VA_t^i &= VA_{t-4}^i * (1 + g_t^i) \\ g_t^i &= \frac{X_t - X_{t-4}}{X_{t-4}} \end{aligned}$$

where VA_t^i indicates value added inf sector i and g_t^i is the annual growth rate in the corresponding benchmark physical indicator X. An estimate of quarterly GDP is then reached by aggregating the sectoral components.

2.2 Production side estimation from monthly variables

A fairly detailed description of the main indicators employed as proxies by the CSO is well documented in the national accounts manuals [Nat, 2007], however a 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 [3] growth, from a small set of monthly indicators. For each sectoral value added. Through this, we attempt to

[3] Henceforth, by GDP we mean either GDPxagri or GDPxoth

reconstruct what the CSO does every quarter. We are able to replicate the exercise fairly successfully for some of the sectors, however for some of the sectors we do not have access to all the monthly indicators used by the CSO.

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 add some monthly indicators that we think might have some impact on the sectoral value added. As an example, we add turnover on the NSE as one of proxy indicators for GDP (Banking and insurance). The methodology essentially relies on bridge equations, developed to link early monthly releases with quarterly GDP growth for each sectoral value added. Section 3 describes the underlying model structure applied in the reconstruction process.

Table 1 Indicators used for quarterly estimates of 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, CPI

Figure 3 Reconstructing sectoral GDP excluding agriculture growth from monthly indicators

These graphs show the reconstruction exercise for the various sub-sectors of GDP. The blue line shows the actual year-on-year growth rate of the sub-sectors of GDP and the green line shows the reconstructed sectoral GDP from the available monthly indicators.

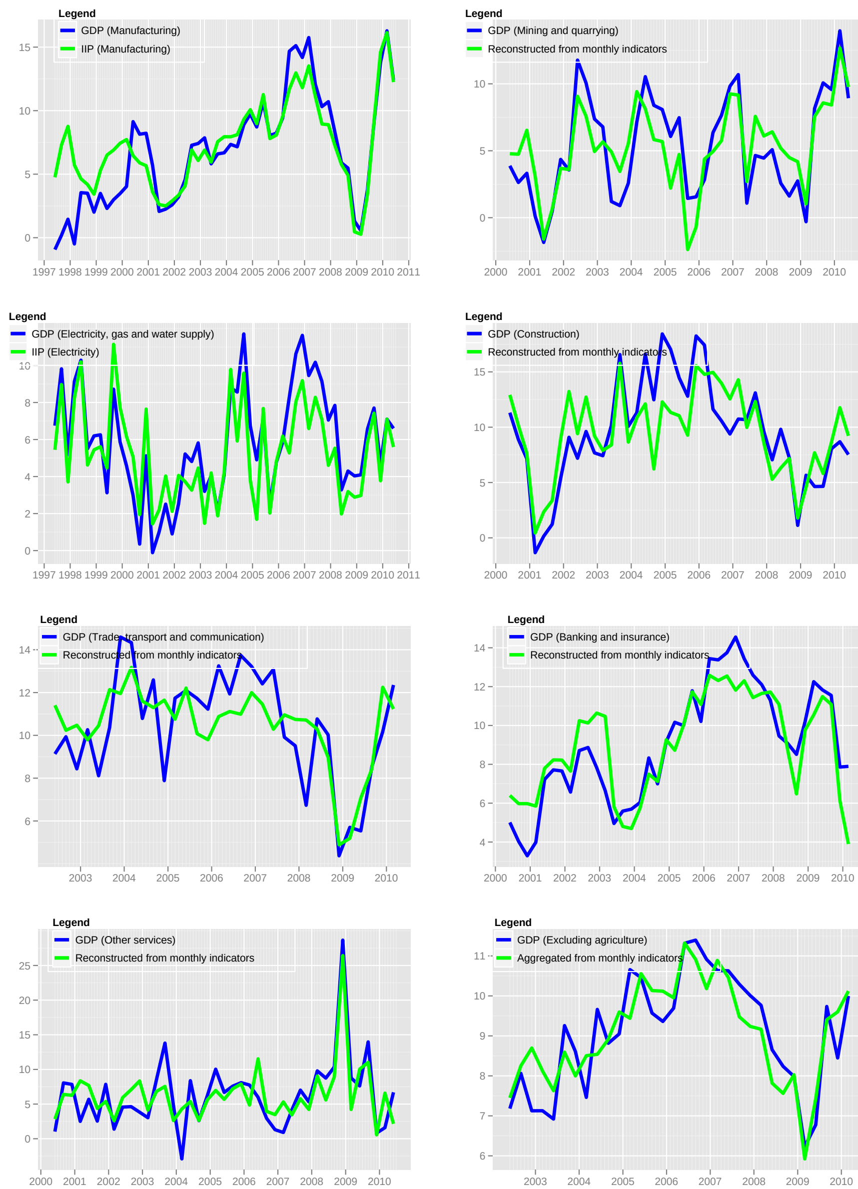


Figure 3 shows the reconstruction exercise for each sub-sector of GDP. Our analysis shows that for some of the sectors, the fit is relatively better than the others. Particularly, the end of sample fit is good. The fit is not good within-sample because we are comparing the monthly indicators with the revised and not the actual GDP estimates. The quarterly growth rates (Q1, Q2 and Q3) are revised with the Q1 (Jan-March) data, thereafter they are revised with the corresponding annual estimates. Our monthly indicators serve as proxies for the initial estimates of GDP. The revised estimates of GDP rely on different data sources. The initial quarterly estimates are based on, for example monthly IIP data but shift to Annual Survey of Industries (ASI) data when annual estimates of GDP are released and the quarterly estimates are revised.

The bottom right panel of Figure 3 shows the reconstructed non-agricultural GDP through aggregation of sectoral monthly indicators. The R square of the model reconstructing GDP growth using GDP excluding both agriculture and services as target is 0.89, while that for the model using only non-agricultural GDP as target is 0.73.

Figure 4 Reconstructing sectoral GDP excluding agriculture and other services growth from monthly indicators

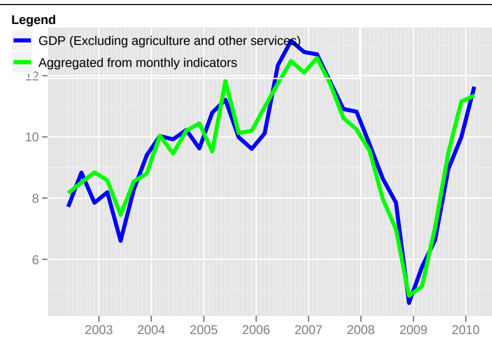


Figure 4 shows the growth rate of reconstructed GDP excluding agriculture and other services using of sectoral monthly indicators.

2.3 Orthogonal 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 i) they 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 is sensitive to sample size and composition. A number of papers [Angelini et al., 2008, Matheson et al., 2007], investigate the forecasting 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 type 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 26th July, 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.[4]

[4]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 utilization, level of capacity utilization (compared to

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 month July is 2nd August, 2010.

The *D&B* Business Optimism Index for India is well known among investors and policymakers. The survey is released a few days after the end of each quarter. The index is arrived at 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 6 questions on business developments over the past and next year.[5]

the average in four quarters), assessment of the production capacity with regard to expected demand in next six months, employment in the company, exports, imports, selling prices, increase in selling prices and profit margin.

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

3 The real-time forecasting exercise

Our out of sample forecasting exercise spans a period of 6 years from 2005Q2 to 2010Q2. The forecasting accuracy of different methods is evaluated under realistic informational assumptions. The forecasting evaluation exercise mimicks as closely as possible the real-time flow of information, by replicating the real time pattern of data availability. The parameters of the model are estimated recursively using only the information available at the time of the 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 31 August 2010 and combine this with the typical data release calendar to re-construct data availability at the end of each month.

We consider two alternative targets: GDP excluding agriculture, GDP excluding agriculture and other services. The forecasting models that we consider are the following:

- benchmarks:
  - Naive model (random walk forecast)
  - AR model
- Autoregressive models with exogenous proxies (ARX), where the indicators are:
  - real activity variables, financial variables, etc.
  - survey variables
  - international activity and survey variables
- *bottom-up* bridge models from sectoral value added

For GDP of a given quarter, we produce a sequence of forecasts in 3 consecutive months prior to the release of the official quarterly GDP. We will label these 3 sequences, as *month-0*, *month-1*, *month-2*, respectively denoting the forecasts 0 months from the GDP release, 1 and 2 month ahead from of it. Starting from the N variable dataset extracted on 31 August 2010 (T), $\Omega_T = \{x_s\}_{s=1}^{T}$, 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 $x_{j,t-h}$, $h \geq 0$ and $j = 1, \cdots, N$ if observations $x_{j,T-h}$ are missing in $\Omega_T = \{x_s\}_{s=1}^{T}$.

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. On Aug. 31 2010, with a delay of two months with respect to end of the reference quarter (31 March 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 Aug.12. On the other hand, the information flow on commercial vehicles production is more timely: on 31Aug. 2010 data up to July was available, having been released on 11 Aug.2010.

Table 2 Data available on 31 Aug. 2010, just ahead of Q2 GDP release

	Mar-10	Apr-10	May-10	Jun-10	Jul-10	Aug-10	Last release
IIP	X	X	X	X			12Aug2010
Cement	X	X	X	X	X		26Aug2010
Steel	X	X	X	X	X		26Aug2010
Coal	X	X	X	X	X		26Aug2010
Railway	X	X	X	X	X		26Aug2010
Ports	X	X	X	X	X		26Aug2010
Tourists	X	X	X	X	X		27Aug2010
Vehicles	X	X	X	X	X		11Aug2010
Electricity	X	X	X	X	X		02Aug2010
Phones	X	X	X	X	X		13Aug2010
Credits	X	X	X	X	X		14Aug2010
Deposits	X	X	X	X	X		14Aug2010
Govexp	X	X	X	X			28Aug2010
BSE	X	X	X	X	X	X	31Aug2010
PMI surveys	X	X	X	X	X		6 Aug2010
US IIP	X	X	X	X	X		5Aug2010

3.1 The models at work

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.

3.1.1 AR+X models

Let us denote growth (year-on-year) in our quarterly target variable as g_t^Q and the vector of k selected monthly indicators, for every AR+X model j, as $x_t^j = (x_{1,t}^j, \cdots, x_{k,t}^j)'$, $\quad t = 1, \cdots, T$. The models are estimated from quarterly aggregates of the monthly data. Predictions of the target GDP series are obtained in two steps. In the first step, the monthly indicators are forecast 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 the resulting quarterly aggregates are used as regressors in the bridge equation to obtain the GDP forecast, with the the following structure:

$$g_t^Q = \mu + \phi_2 g_{t-1}^Q + \sum_{i=1}^{k} \beta_i^j(L) x_{i,t}^{j,Q} + \epsilon_t^{j,Q} \tag{1}$$

where μ is an intercept parameter and $\beta_i^j(L)$ denotes a lag polynomial.

3.1.2 Bridge bottom up

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

$$VA_r^Q, t = \mu + \phi_2 VA_{r,t-1}^Q + \sum_{i=1}^{k} \beta_i^r(L) x_{i,t}^{r,Q} + \epsilon_t^{j,Q} \tag{2}$$

where $r = 1, \cdots, R$ are the R sectors making up the GDP.

Finally, the growth rate of target variable, g_t^Q, is obtained by a second stage regression using the predicted sectoral value added growth rates VA^Q:

$$g_t^Q = \delta + \sum_{r=1}^{R} \gamma_n \hat{VA}_r^Q + \epsilon_t \tag{3}$$

This approach of obtaining over all growth rate by aggregating the monthly proxies for various sub sectors of GDP is referred to as the *bottom up approach*.

3.2 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 2005 Q2 to 2010 Q2. We focus separately on the results relating to the two target variables. As to our first target, GDP excluding agriculture and other services (GDPxoth), the results are reported in Table 3.

What follows next, is an attempt to summarize how information flows over time and across sectors help us to gauge the growth rate in GDPxoth. 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 1.3 percentage points. As we start adding information (moving from month- 2 to month-0) the precision of the forecasts improves: some of the models using a single indicator (AR+X models) show a sizeable reduction in the RMSE (of approximately 16%). Among these, the one based on the Index of Industrial production, and the PMI global indicator perform the best. Interestingly, the quarterly Indian survey data, while improving slightly the RMSE compared to the benchmark models, do not seem to contain additional predictive content with respect to the models containing more timely real indicators. Finally, we find that, for the 3 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 26% from 0.88 at month 2, to 0.65 at month-0. In line with our in sample results from Section 2.2, we find that also out sample the 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 financial sector (deposits, loans, BSE and WPI) does not lead to a significant reduction of the RMSE of our forecasts. 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.

Table 4 presents RMSE of alternative models when the target variable is non-agricultural GDP. For this alternative target too, bridge models outperform models relying on quarterly variables. However the relative ranking of these models changes. In particular, the simpler bridge model relying only on IIP manufacturing performs better than the more complex sub-sectoral bridge model. Looking more closely into the results 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 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 particulary 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 behavior. Again, like in the case of GDPXoth, the Indian survey data, while improving the RMSE compared to the benchmark models, do not seem to contain additional predictive content with respect to the models containing more timely real indicators.

Figure 5 shows the three consecutive forecasts from the bridge bottom-up, obtained in pseudo real time, vis-a-vis the (single) forecast from the AR model which can be obained as soon as a new GDP figure is released. We find that bridge model is able to track the movements in the GDPoxth growth more accurately relative to the AR model. Also, it is evident from the figure that as expected, the forecast at zero month from GDP release provides a better fit as we have access to more information about the corresponding quarter. However in terms of pseudo real time forecast of GDPxagri growth, performance of the bridge model is poor.

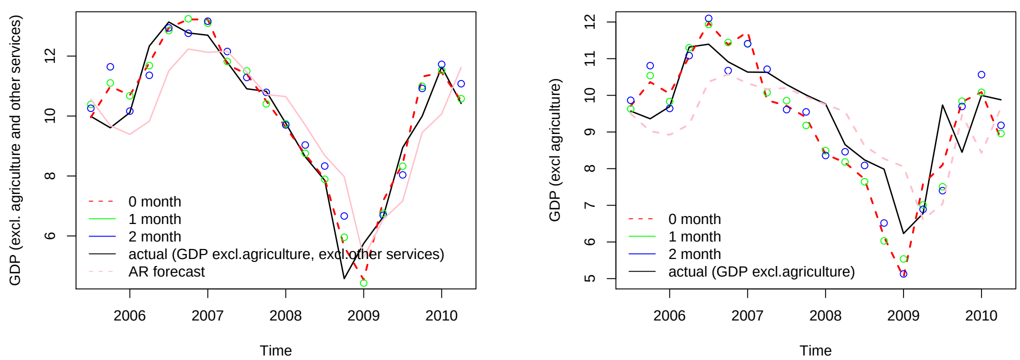
Table 3 RMSE of alternative models for target: GDP excluding agriculture and other services

Model	Sector	Forecast	Nowcast		
			Month 0	Month 1	Month 2
Naive		1.31			
AR		1.27			
Bridge bottom up			0.65	0.69	0.88
	Manufacturing		1.27	1.53	1.57
	Electricity		1.66	1.83	1.82
	Trade and transport		1.53	1.53	1.56
	Finance		1.42	1.42	1.43
	Construction		2.56	2.56	2.60
	Mining		2.18	2.35	2.38
AR+X models					
IIP Manufacturing			0.75	0.77	0.89
Financial situation		0.95			
Availability of finance		1.05			
Inventory of finished products		3.63			
Assessment of production capacity for 6 months		3.47			
PMI Global			1.08	1.08	1.09
D&B business survey		3.87			
US (IIP)			1.252	1.258	1.26

Table 4 RMSE of alternative models for target: non-agricultural GDP

Model	Sector	Forecast	Nowcast		
			Month 0	Month 1	Month 2
Naive		1.04			
AR		1.05			
Bridge bottom up			0.93	0.95	0.92
	Manufacturing		1.270	1.534	1.57
	Electricity		1.657	1.830	1.82
	Trade and transport		1.532	1.532	1.56
	Finance		1.417	1.525	1.43
	Construction		2.557	2.557	2.60
	Mining		2.180	2.346	2.38
	Other services		3.11	4.13	4.29
AR+X					
IIP Manufacturing			0.77	0.77	0.95
Financial situation				0.73	
Availability of finance		0.78			
Inventory of finished products		2.05			
Assessment of production capacity for 6 months		1.70			
PMI Global			0.88	0.88	0.89
Bradbury business survey		2.30			
US IIP			0.92	0.927	0.93

Figure 5 AR forecasts vs bridge bottom up (0-1-2 months ahead of GDP release)



4 Conclusions

This paper applies bridge models to forecast short-term GDP growth in India. In essence, 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 aggregated year-on-year growth rate 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, 1 month and few days before the actual GDP release. The results of the nowcasting exercise show that bridge models significantly outperform the benchmark AR and naive models.

Finally, we investigate for the first time the effective usefulness of Indian survey data in nowcasting GDP. Our results suggest that the survey data available for India is not enhancing the predictive accuracy of our nowcasts. In particular, the Reserve Bank of India business survey, given its quarterly nature, and its small time advantage with respect to the GDP release, is find to be of little use to nowcast GDP, as hard data are already available covering most of the reference quarter. This finding applies not only for the survey responses regarding the current quarter assessment, but also for those referring to expected movements in the following quarter.

References

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

Report of the Technical Advisory Group on Development of Leading Indicators for Indian Economy. Technical report, Reserve Bank of India, 2007.

F. Altissimo, A. Bassanetti, R. Cristadoro, M. Forni, M. Hallin, M. Lippi, L. Reichlin, and G. Veronese. A real time coincident indicator for the euro area business cycle. *CEPR Discussion Paper No. 3108.*, 2001.

F. Altissimo, R. Cristadoro, M. Forni, M. Lippi, and G. Veronese. New eurocoin: Tracking growth in real-time. *Review of Economics and Statistics, forthcoming*, 2009.

T.W. Anderson and TW Anderson. *An introduction to multivariate statistical analysis.* Wiley New York, 1958.

E. Angelini, G. Camba-Mendez, D. Giannone, G. Runstler, and L. Reichlin. Short-term forecasts of Euro Area GDP growth. *European Central Bank Working Paper No.949*, 2008.

A. Baffigi, R. Golinelli, and G. Parigi. Bridge models to forecast the euro area GDP. *International Journal of Forecasting*, 20, Issue 3:447–460, 2003.

J. Bai. Inferential theory for factor models of large dimensions. *Econometrica*, 71(1):135–171, 2003.

J. Bai and S. Ng. Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221, 2002.

M. Banbura and D. Giannone. Nowcasting. *ECARES working paper 2010âĂŘ021*, 2010.

K. Barhoumi, Benk S., Cristadoro R., Den Reijer A., Jakaitene A., Jelonek P., Rua A., Runstler G., Ruth K., and Van Nieuwenhuyze C. Short term forecasting of GDP using large datasets: a pseudo-realtime forecast evaluation. *Occasional Paper Series, European Central Bank*, 2008.

M. Baxter and R.G. King. Measuring business cycles: approximate band-pass filters for economic time series. *Review of Economics and Statistics*, 81(4):575–593, 1999.

B.S. Bernanke and J. Boivin. Monetary policy in a data-rich environment* 1. *Journal of Monetary Economics*, 50(3):525–546, 2003.

P.J. Brockwell and R.A. Davis. *Time series: theory and methods*. Springer Verlag, 2009.

A.F. Burns and W.C. Mitchell. Measuring business cycles. *NBER Books*, 1946.

G.C. Chow and A. Lin. Best linear unbiased interpolation, distribution, and extrapolation of time series by related series. *The Review of Economics and Statistics*, 53(4):372–375, 1971.

L.J. Christiano and T.J. Fitzgerald. The Band Pass Filter*. *International Economic Review*, 44 (2):435–465, 2003.

M.D. Evans. Where are we now? real-time estimates of the macro economy. *NBER Working Paper*, 2005.

M. Forni, M. Hallin, M. Lippi, and L. Reichlin. The generalized dynamic-factor model: Identification and estimation. *Review of Economics and Statistics*, 82(4):540–554, 2000.

M. Forni, M. Hallin, M. Lippi, and L. Reichlin. Coincident and leading indicators for the euro area. *The Economic Journal*, 111(471):62–85, 2001.

M. Forni, M. Hallin, M. Lippi, and L. Reichlin. The generalized dynamic factor model. *Journal of the American Statistical Association*, 100(471):830–840, 2005.

D. Giannone, L. Reichlin, and L. Sala. Tracking Greenspan: systematic and unsystematic monetary policy revisited. *CEPR Discussion Paper No. 3550.*

D. Giannone, L. Reichlin, and S. Simonelli. Nowcasting Euro Area Economic Activity in Real-Time: The Role of Confidence Indicators. *ECARES working paper 2009-021*, 2009.

T. Matheson, J. Mitchell, and B. Silverstone. Nowcasting and predicting data revisions in real time using qualitative survey data. *Reserve Bank of New Zealand, Discussion Paper 2007/02*, 2007.

Reserve Bank of India. Report of the Working Group on Surveys. Technical report, Reserve Bank of India, 2009.

T.N. Srinivasan. India's Statistical System: Critiquing the Report of the National Statistical Commission. *Economic and Political Weekly*, 38(4):303–306, 2003.

J.H. Stock and M.W. Watson. New indexes of coincident and leading economic indicators. *NBER macroeconomics annual*, pages 351–394, 1989.

J.H. Stock and M.W. Watson. Business cycle fluctuations in US macroeconomic time series. *Handbook of Macroeconomics*, 1:3–64, 1999.

J.H. Stock and M.W. Watson. Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97(460):1167–1179, 2002a.

J.H. Stock and M.W. Watson. Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, 20(2):147–162, 2002b.

J. Valle e Azevedo, S.J. Koopman, and A. Rua. Tracking the Business Cycle of the Euro Area. *Journal of Business and Economic Statistics*, 24(3):278–290, 2006.

V. Zarnowitz and A. Ozyildirim. Time series decomposition and measurement of business cycles, trends and growth cycles. *Journal of Monetary Economics*, 53(7):1717–1739, 2006.