NowCasting Indian GDP Growth: Dynamic Factor Model

HUL715: Time Series Econometrics

S S K Pavan
2020HES7045

November 29, 2021

Nowcasting

prediction of the present, the very near future and the very recent past[1]

- Key statistics on the state of economy available after lag
- Aim: Obtain early estimate by exploiting information which is published early

[1] Banbura, Giannone, and Reichlin (2013)

Nowcasting

prediction of the present, the very near future and the very recent past[1]

- Key statistics on the state of economy available after lag
- Aim: Obtain early estimate by exploiting information which is published early
- Diverse indicators show varied trends in economic activity
- Nowcasting literature builds formal statistical models to assimilate these varying trends

[1] Banbura, Giannone, and Reichlin (2013)

Motivation

Mint Newspaper Plain facts section

- 1st Nov: Coming this week: India PMI data; US, UK monetary policy meets
- 25th Oct: What to watch: COP26 meeting, auto & pharma earnings
- 18th Oct: Coming up this week: MPC minutes, FMCG results, UK inflation
- 11th Oct: Numbers to watch: Inflation, Q2 earnings in India; IMF global outlook
- 27th Sept: Numbers to watch: Core sector data, PMI in India; GDP in the US, UK

Challenges

- Large number of variables sampled at wide range of frequencies
- Data released in non-synchronous manner, with different delays

Challenges

- Large number of variables sampled at wide range of frequencies
- Data released in non-synchronous manner, with different delays

- Sub-challenges:
  - Choice of variables
    - Lead
    - Lag
    - Contemporaneous
  - Signal Extraction

Methods

- Bridge Equations
- Mixed Data Sampling
- Mixed Frequency VAR
- **Dynamic Factor Models**
- Bayesian VAR and Bayesian DFM

Literature - I

- Until start of 2000s, bridge equations were popular for short term forecasting (Baffigi, Golinelli, and Parigi, 2004)
- Giannone, Reichlin, and Small (2008) devised formal statistical framework for use in nowcasting.

Literature - I

- Until start of 2000s, bridge equations were popular for short term forecasting (Baffigi, Golinelli, and Parigi, 2004)
- Giannone, Reichlin, and Small (2008) devised formal statistical framework for use in nowcasting.
- Factor analysis has long history in economics:
  - Gweke(1977) captured the co-movements in economic time series
  - Sargent and Sims (1977), Stock and Watson (1989, 1991, 1993) used dynamic factors to identify business cycles

Literature-II

- Nearly every central bank has a nowcasting model in place
- Two new papers using Bayesian methods:
  - Dıaz, Drechsel, Petrella 2021 use Bayesian Dynamic Factor model with novel specifications
  - Cimadomo, Giannone, Lenza, Monti, Sokol 2021 use large Bayesian Vector Autoregressions with a uge dataset

Literature-II

- Nearly every central bank has a nowcasting model in place
- Two new papers using Bayesian methods:
  - Dıaz, Drechsel, Petrella 2021 use Bayesian Dynamic Factor model with novel specifications
  - Cimadomo, Giannone, Lenza, Monti, Sokol 2021 use large Bayesian Vector Autoregressions with a uge dataset
- Indian context:
  - RBI 2007 Composite Index of Coincidental Indicators
  - Bragoli and Fosten (2018) use DFM
  - Rabobank and CEIC publish nowcasts of Indian GDP using proprietary models

Variable Selection: Using Literature

- Step 1: Based on literature review of High-frequency indicators of Indian economy, authors choose **28 potential indicators**

Industry	Income	Employment	Services	External Sector	Prices	Misc. Activity	Credit
IIP - Core	IIP-Cons goods	Agri wages	**Air cargo**	**non-oil, gold imports**	CPI	**Govt Tax receipts**	**commercial credit**
Oil cons	**Auto sales**		**Rail Freight**	**Exports**			**Sensex**
Power demand			Air passengers	**tourist arrival**			Treasury yield
Steel prod			Ral passenger	**NEER**[2]			
Cement prod				Forex Reserves			
				CRB index			
				crude prices			
				OECD IP			
				CBOE VIX			

- Step 2: Choose subset which has significant correlation with GDP at $t-1$, t and $t+1$
- Step 3: Check if the chosen indicators can successfully identify turning points in GDP growth rate

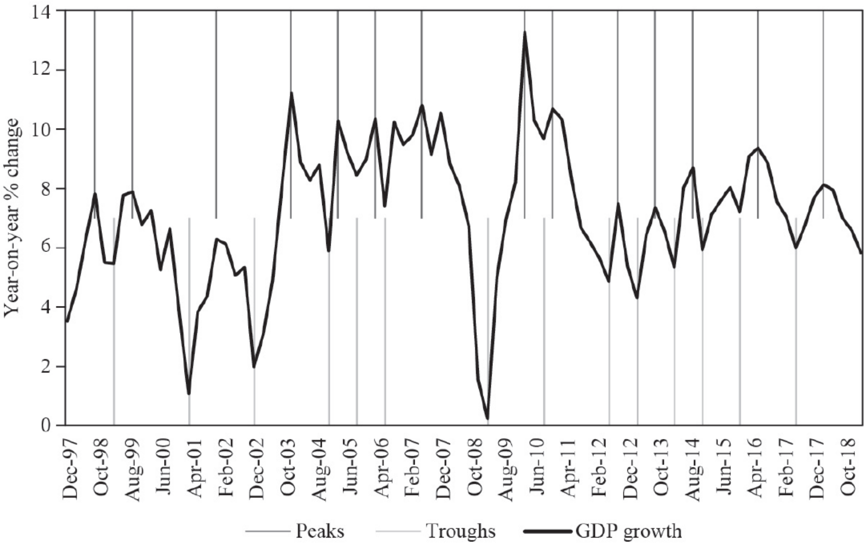
[2] Nominal Effective Exchange rate

Variable Selection: Dynamic Correlation

Indicators	$GDP_{(t-1)}$	$GDP_{(t)}$	$GDP_{(t+1)}$	Indicator Type
Air cargo	0.17	0.37*	0.27*	C
Auto total	0.17	0.25*	0.24*	C
Bank credit	0.11	0.08	−0.09	X
Exports	0.31*	0.35*	−0.02	C
Foreign tourist	0.19	0.36*	0.12	C
Government receipts	0.21*	0.10	−0.01	L−
IIP consumer goods	0.20*	0.54*	−0.01	C
IIP core	0.25*	0.20*	0.28*	L+
NEER	0.14	0.24*	0.15	C
NONG imports	0.33*	0.19	−0.02	L−
Rail freight	0.20*	0.17	0.15	L−
Sensex	0.25*	0.52*	0.23*	C

Figure: summary of selected variables

Turning Points



LASSO

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

- n is the number of turning points in from Q1 2003 to Q1 2019
- y_i is the quarter-on-quarter annualized growth rate of GDP
- x_{ij} is the jth high freq indicator at ith turning point
- λ is the tuning parameter

LASSO

Indicator	Turning-Point Correlation	Lasso
IIP consumer goods	0.68*	Y
Auto total	0.55*	Y
NEER	0.55*	Y
IIP core	0.58*	Y
Rail freight	0.24*	Y
Sensex	0.65*	Y
Exports	0.65*	Y
Air cargo	0.57*	N
NONG imports	0.46*	N
Bank credit	0.11	N
Government receipts	0.32*	N
Foreign tourist	0.57*	N

Figure: Summary of Selected Variables Using Turning-Point Correlations and Lasso

samples

- CEII - 6 Domestic economic activity indicators
- CEII - 9 trade and service sector
- CEII - 12 financial sector

Basic DFM I

- Variables are log differenced, demeaned and standardised
- Model:

$$
\begin{aligned}
Y_t &= \Lambda f_t + E_t, \qquad E_t \quad i.i.d\ N(0, R) \\
f_t &= \beta f_{t-1} + u_t, \qquad u_t \quad i.i.d\ N(0, q)
\end{aligned}
$$

- Y_t contains only monthly indicators from a balanced panel
- Assumption: f_t and v_{it} are mutually uncorrelated at all lags and leads

Basic DFM I

- Variables are log differenced, demeaned and standardised
- Model:

$$
\begin{aligned}
Y_t &= \Lambda f_t + E_t, \qquad E_t \quad i.i.d\ N(0, R) \\
f_t &= \beta f_{t-1} + u_t, \qquad u_t \quad i.i.d\ N(0, q)
\end{aligned}
$$

- Y_t contains only monthly indicators from a balanced panel
- Assumption: f_t and v_{it} are mutually uncorrelated at all lags and leads
- 2-Step procedure:
  - Parameters are estimated using PCA
  - Factor f_t is estimated by applying Kalman-smoother to the entire information set

PCA I

- Suppose $\sum_{yy}$ is the var-covariance matrix of the observed variables.

$$Y_t = \Lambda F_t + E_t, \qquad E_t \quad i.i.d \; N(0, R)$$

- The factors F_t (q x 1) are modelled as mean 0, variance 1 and are orthogonal to each other i.e $F_t F_t' = 0$
- take the variance of Y_t,

$$\begin{aligned}\sum_{yy} &= \Lambda F_t F_t' \Lambda' + R \\ &= \Lambda \Lambda' + R\end{aligned}$$

PCA- II

- $\sum_{yy}$ is a positive definite matrix
- It can be decomposed as follows:

$$\sum_{yy} = eDe' = eD^{1/2}D^{1/2\prime}e'$$

- e is the matrix of eigen vectors (p) and D is a diagonal matrix of eigen values

PCA- II

- $\sum_{yy}$ is a positive definite matrix
- It can be decomposed as follows:

$$\sum_{yy} = eDe' = eD^{1/2}D^{1/2\prime}e'$$

- e is the matrix of eigen vectors (p) and D is a diagonal matrix of eigen values
- estimate:

$$S_{yy} \approx \lambda_1 e_1 e_1' + \ldots + \lambda_q e_q e_q' = \hat{\Lambda}\hat{\Lambda}'$$

- remaining $\lambda_{q+1} \ldots \lambda_p$ are dropped
- estimate R as the residual

$$R = S_{yy} - \hat{\Lambda}\hat{\Lambda}'$$

State-Space

- Original model:

$$
\begin{aligned}
Y_t = \Lambda f_t + E_t, &\quad E_t \quad i.i.d\ N(0, R) \\
f_t = \beta f_{t-1} + u_t, &\quad u_t \quad i.i.d\ N(0, q)
\end{aligned}
$$

- We have estimates of Λ and R. We assumed q $= 1$
- using the kalman smoother to the entire information set, the factors can be estimated.
- Estimation done through EM algorithm. Software to implement available in R through the MARSS package and in python through Statsmodels

Factors

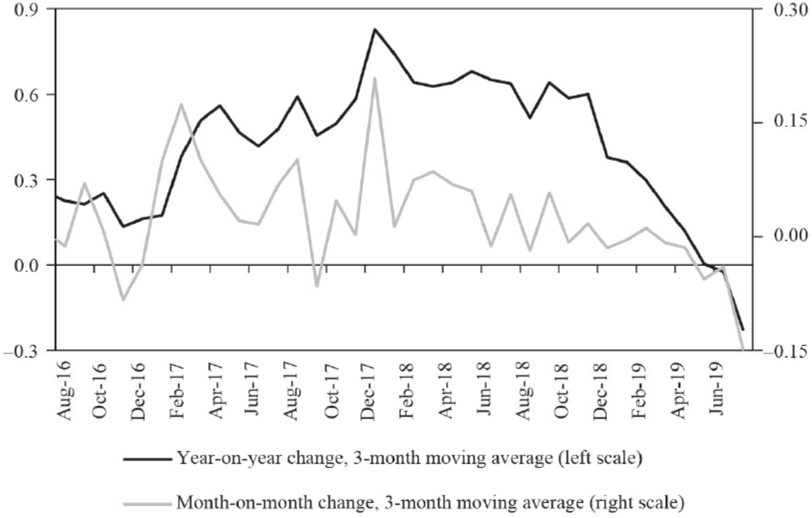


Figure: Growth in CEII-6

factors

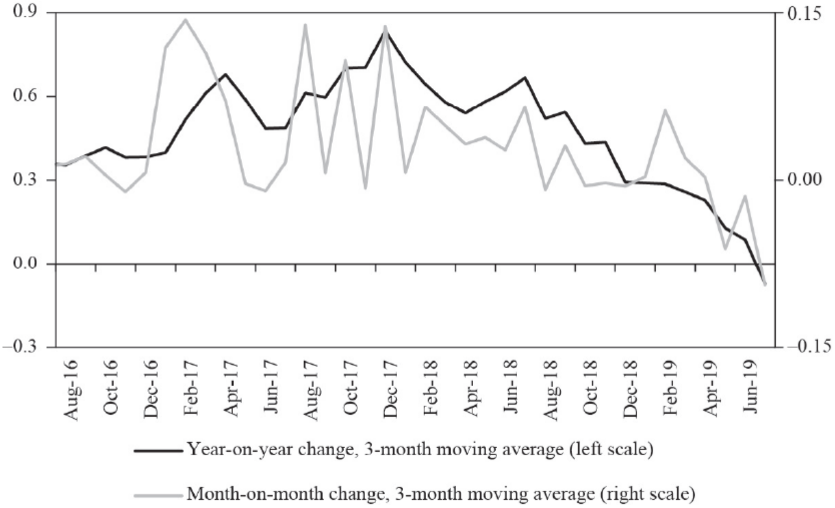


Figure: Growth in CEII-9

factors

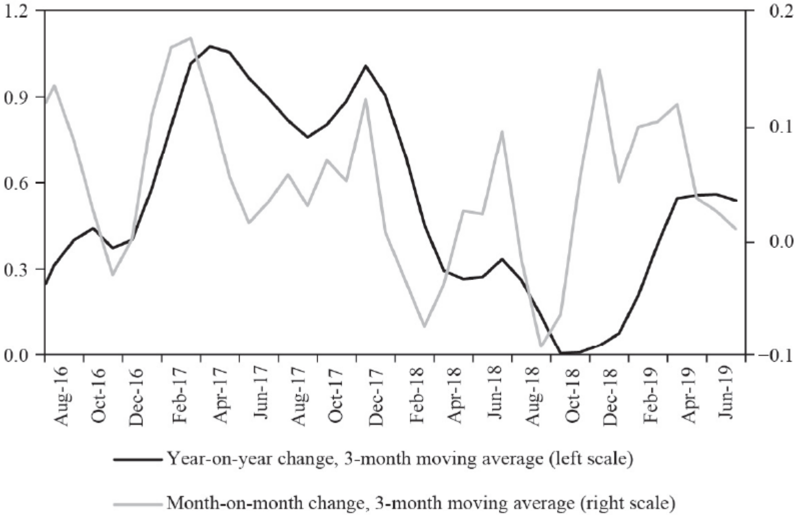


Figure: Growth in CEII-12

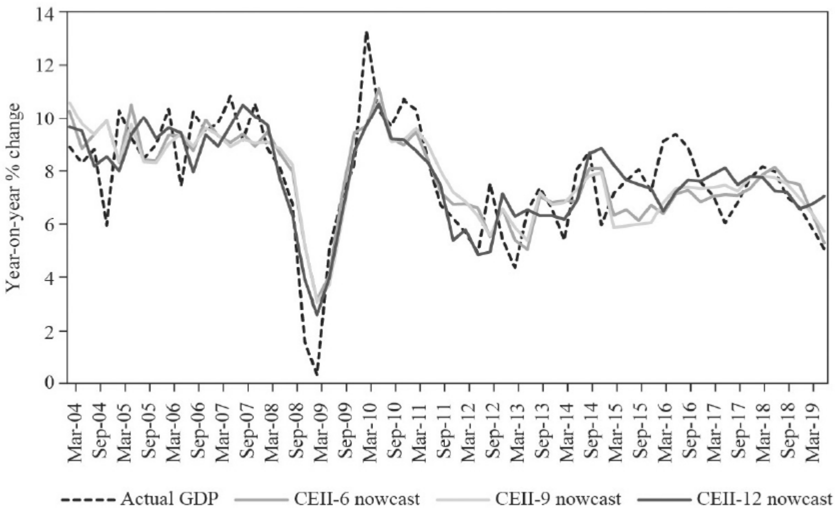
Nowcasting

Table 3. **Nowcasting Model Estimates**

	GDP (Y-o-Y)		
Dependent Variable	**Model 1 (6 indicators)**	**Model 2 (9 indicators)**	**Model 3 (12 indicators)**
Constant	3.02	3.64	4.60
	(0.69)	(0.73)	(0.73)
CEII (Y-o-Y)	2.94	3.08	1.47
	(0.64)	(0.70)	(0.27)
GDP (Y-o-Y), Lag 1	0.38	0.28	0.27
	(0.10)	(0.12)	(0.10)
Model Diagnostics			
Adjusted R-squared	0.54	0.53	0.59
B-G Ser. Corr. LM Test	0.19	0.11	0.55
Out-of-Sample RMSE	0.61	0.65	0.98

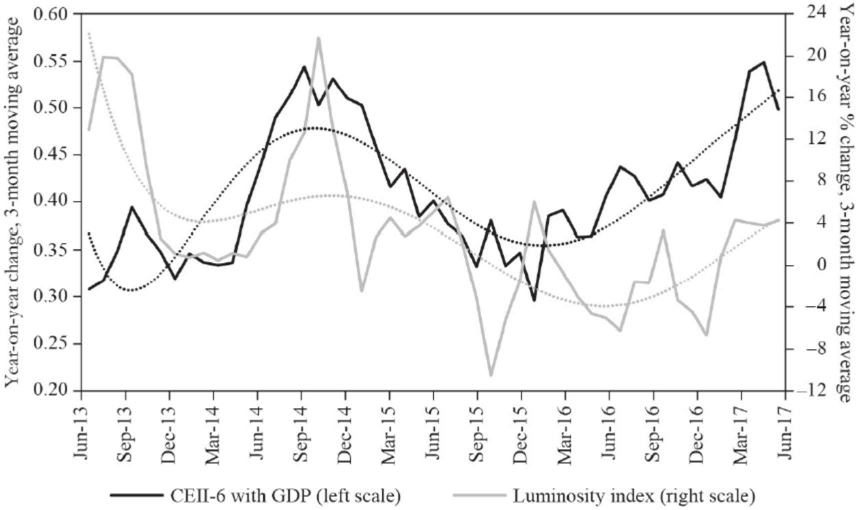
B-G Ser. Corr. LM = Breusch–Godfrey serial correlation Lagrange multiplier, CEII = Coincident Economic Indicators for India, GDP = gross domestic product, RMSE = root mean squared error, Y-o-Y = year-on-year.
Notes: Estimates are based on the sample period Q1 2004 to Q1 2019. The total number of observations is 62. All coefficient estimates are significant at 1%. Standard errors are reported in parentheses. The B-G test is used to assess for serial correlation in errors up to 12 lags. Out-of-sample RMSE pertains to Q1 2017 to Q1 2019.

Nowcasting



14
12
10
8
6
4
2
0
Year-on-year % change
Mar-04
Sep-04
Mar-05
Sep-05
Mar-06
Sep-06
Mar-07
Sep-07
Mar-08
Sep-08
Mar-09
Sep-09
Mar-10
Sep-10
Mar-11
Sep-11
Mar-12
Sep-12
Mar-13
Sep-13
Mar-14
Sep-14
Mar-15
Sep-15
Mar-16
Sep-16
Mar-17
Sep-17
Mar-18
Sep-18
Mar-19
Actual GDP
CEII-6 nowcast
CEII-9 nowcast
CEII-12 nowcast

validation



Year-on-year change, 3-month moving average
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
Year-on-year % change, 3-month moving average
24
20
16
12
8
4
0
−4
−8
−12
Jun-13
Sep-13
Dec-13
Mar-14
Jun-14
Sep-14
Dec-14
Mar-15
Jun-15
Sep-15
Dec-15
Mar-16
Jun-16
Sep-16
Dec-16
Mar-17
Jun-17
CEII-6 with GDP (left scale)
Luminosity index (right scale)

Conclusion

- What did we do?
  - Built a single indicator which extracts common trend underlying high frequency indicators
  - Forecasted GDP using common factor