A Multi-Factor GDP Nowcast Model for India

Kaustubh Kaustubh* Abhishek Ranjan†

Abstract

Post-pandemic challenges have necessitated a re-evaluation of nowcasting models for Indian GDP. A significant issue is the availability of diverse high-frequency indicators (HFIs) with distinct lead-lag relationships to GDP. We have used a mixed-frequency multi-factor vector autoregressive (VAR) model to incorporate various HFIs in our nowcasting model. The data in the model, include nominal HFIs, survey-based HFIs, HFIs related to the labor market, and HFIs related to real economic activity. Current nowcasting models for India focus solely on overall GDP nowcasts, neglecting the individual contributions of each HFI. To address this, we calculate the impact of each HFI on GDP nowcast revisions with new data releases. Another challenge arises from COVID-19 disruptions, which have caused outlier values in HFIs and distorted model parameters. We introduce the Oxford Stringency Index and a novel data transformation to address this challenge, which reduces the sensitivity of models to large shocks and enables them to handle unexpected events without overreacting. This transformation can add significant value to the forecasting literature involving outliers.

Keywords: Nowcasting, News, Mixed-Frequency Dynamic Factor Model

JEL: E52, C53, C33.

*Manager Research, Strategic Research Unit, RBI, Email: *kaustubh@rbi.org.in* .

†Manager Research, Strategic Research Unit, RBI, Email: *abhishekranjan@rbi.org.in.* All views and opinions expressed reflect those of the authors and not necessarily of the Reserve Bank of India.

1 Introduction

Macroeconomic policymaking relies on forward-looking assessments of the economy, yet official estimates of key macroeconomic indicators such as GDP often come with a lag. To address this gap, central banks have increasingly relied on nowcasting models, which provide timely, forward-looking insights into vital macroeconomic indicators like GDP. Nowcasting involves estimating the current or near-future state of a lower-frequency target variable using available high-frequency indicators (HFIs) prior to the release of official estimates of the target variable. For example, we can use nowcasting to estimate the current and near-future state of India's GDP, which is released at quarterly frequency with an eight-week lag. The HFIs, available at monthly, weekly, or daily frequencies, contain relevant information about GDP ahead of official GDP numbers. Thus, nowcasting is distinct from forecasting as it emphasizes measuring the target variable's current status rather than predicting its future state.

Following its economic liberalization in 1991, India, which was largely a closed economy before then, has fully integrated into the global economy. Global events, such as the 1997-Asian Financial Crisis, the 2007-08 Global Financial Crisis, and other events, including COVID-19, have a visible impact on the Indian economy. Since reforming and opening up, the government has implemented a series of reforms to restructure and upgrade the economy.

India's reforms have positioned the country as the 5^{th} largest in terms of GDP, with projections to rise to 3^{rd} place by *2030*. Maintaining macroeconomic stability is crucial for maintaining this growth trajectory. In this regard, an improved nowcast of Indian GDP would offer valuable insights into the current economic landscape, aiding in macroeconomic surveillance for policymakers. Furthermore, given the significant impact of GDP announcements on financial markets, precise nowcasting can help align market expectations with actual economic performance. Overall, improving India's GDP nowcast performance may help to promote economic stability, guide decision-making processes, and facilitate efficient resource allocation.

Nowcasting GDP for emerging markets, including India, is extremely challenging due to the relatively short time span of available macroeconomic data, missing observations during the sample periods, variances in data quality, and the swiftness of continuous structural change. Over the last decade, there has been a notable focus on GDP nowcasting in India, aiming to provide accurate estimates for the current and upcoming quarters. Bhattacharya et al. (2011); Dahlhaus et al. (2017); Bragoli and Fosten (2018); Iyer and Gupta (2019a,b); Bhattacharya et al. (2019); Bhadury et al. (2021); Bhattacharya et al. (2023); Ghosh and Ranjan (2023) are notable nowcasting attempts for Indian GDP. Bragoli and Fosten (2018) compare the nowcasts from the first release of GDP to those from the final release to assess the differences in their predictability. Furthermore, the model uses nominal and international series to proxy the missing employment and service sector variables in India, and it finds that predictability of GDP nowcast increases for the final estimates. The informal sector's contribution to the GDP is relatively large, and this contribution is difficult to quantify. To tackle this challenge, it is important to look for survey data (Bhattacharya et al. (2011)) or other proxy series such as nominal and international series as in Bragoli and Fosten (2018), or rainfall deviation as in Iyer and Gupta (2019b). Iyer and Gupta (2019b), show that rainfall has a high predictive content for India as a large proportion of the labour force is involved in agriculture and allied activities. Iyer and Gupta (2019a) use a Bayesian VAR approach to forecast economic growth in India on a quarterly basis, and finds that Bayesian VAR models are able to capture the dynamics of Indian growth well. However, there are several challenges that have emerged in the post-pandemic period that require a redesign of the existing nowcast models for nowcasting Indian GDP.

To begin with, a range of high-frequency indicators (HFIs) providing relevant information on India's GDP are now accessible. This data includes hard statistics on real economic activities, such as the Index of Industrial Production (IIP), survey-based data, such as the Purchasing Managers' Index (PMI), and monthly labor market data, including rural and urban unemployment rates. Different variables show different lead-lag relationships with GDP. For instance, the labor market is generally moving with a lag in economic activity, and survey-based data are expected to be the leading indicators Antolin-Diaz et al. (2024). Thus, a single-factor approach to nowcasting Indian GDP (Bhadury et al. (2021)), or multi-factor approach (Bhattacharya et al. (2023)) in which all HFIs load on all factors, may not be the best approach. To account for potential heterogeneity in the lead-lag relationship of different kinds of HFIs, we propose a mixed-frequency multi-factor VAR model that includes a nominal factor to summarize price information and cost pressures, a soft factor to summarize survey data, and a labor factor to summarize labor market information in addition to real economic activity factor which summarizes information pertaining real economic activity. Our model captures the lead-lag relationships between factors and the target variables by considering both the contemporaneous values of factors and their lags.

The COVID pandemic's disruptions to the data-generating process, which result in outlier values in both target variables and HFIs, present the next challenge for nowcasting models in forecasting Indian GDP in recent years. The outliers distort the model's estimated parameters due to extreme observations, resulting in forecast errors in normal times. To counter this, we have used the Oxford Stringency Index for India as an additional HFI. The Oxford Stringency Index is an index quantifying the extent of lockdown restrictions imposed by different countries at different points in time during the pandemic. Due to lockdown restrictions, we used the stringency index as an additional HFI to ameliorate distortions in the parameter estimates for different HFIs. However, introducing the stringency index is not sufficient to reduce the model parameters' sensitivity caused by the pandemic. Alternatively, to reduce the sensitivity of the model parameters to such shocks, we transform the data using the Sigmoid function[1]. Sigmoid transformation is one of the earliest transformation to be used as an activation function in the neural network literature, and helps in capturing the non-linearity present in the data. The transformation suppresses outliers to zero (one) for large negative (positive) x, while maintaining the linear properties near $x = 0$, and is often used in the forecasting literature (Özkan (2013); Jahn (2020)). In other word, the transformation reduces the sensitivity of data to outliers.

The other challenge that earlier nowcasting models for the Indian economy faced was that they focused only on the final aggregate GDP nowcasts. However, in the case of the revision of nowcasts after the release of new data on HFIs, policymakers are now interested in a more granular analysis and want to know the precise sectors or HFIs that are causing the revision. To our knowledge, no study has systematically identified and quantified the impact of data revisions and new releases from each HFI on India's GDP nowcasting. In this paper, we quantify the impact of any new data release or revision in the pre-released data on GDP nowcasts, and we decompose the impact of data releases on the nowcasts to each HFI. This allows us to precisely identify which HFIs are causing the revisions in the nowcast.[2]

The recent debate to find suitable deflators for estimating India's real GDP growth adds an additional dimension to the existing nowcasting models of Indian GDP. According to the Crisil (2017), non-alignment of manufacturing GDP and IIP is due to divergent WPI and CPI series. Sengupta (2022), has argued that due to the single deflator method used to calculate real GDP in India, in cases of significant divergence between the WPI and CPI, there is divergence between real GDP growth and predicted GDP growth using the HFIs of real economic activity. Therefore, we have incorporated an extra nominal factor into our multi-factor VAR model to consider the potential influence of prices and supply chain pressures on GDP forecasts. We load the nominal factor on the ratio of CPI to WPI in order to capture the divergence between the CPI and WPI, while HFIs, such as the IMF commodity price index and the Baltic dry index, indicate global supply chain pressures.

This paper joins the growing literature on evaluating the relative success of nowcasting models in India. We contribute to the literature in a number of ways. Firstly, in the literature on nowcasting Indian GDP, we decompose the GDP nowcast using blocks of real, nominal, survey, and labor data and examine the impact of new data releases on GDP. Secondly, on the methodological front, we use a sigmoid transformation to reduce the model parameters' sensitivity to large shocks, using a frequentist approach. Our nowcasting model is flexible enough to be adapted for predicting GDP in other emerging market economies facing challenges due to economic shocks like COVID-19.

The rest of the paper is organised as follows: Section 2 deals with the literature review,

[1] $Sigmoid(x) = \frac{1}{1+exp(-x)}$

[2] The decomposition of the impact of data releases on the nowcasts to each HFI is only possible for the multi-factor model with untransformed data, and not in the case of model with transformed data.

and Section 3 focuses on the literature gap and the contributions of the study. Section 4 contains the methodology, and Section 5 discusses the results. Finally, Section 6 provides conclusion of the study.

2 Literature Review

The nowcasting literature has gained traction in the past two decades, addressing the challenge of timely estimation of key economic variables like GDP, which is essential for forward-looking policymaking. Techniques such as dynamic factor models and Bayesian vector autoregressions have emerged as significant tools for nowcasting, emphasizing the importance of early data exploitation for improving forecast accuracy. Another strand of nowcasting literature using machine learning methods Richardson et al. (2021); Ghosh and Ranjan (2023); Zhang et al. (2023) has gained traction in the recent past.

Dynamic factor models have been utilised to construct indices of economic indicators, representing a model-based aggregation approach. Stock and Watson (1989) initially pioneered the application of factor models for developing business cycle indices. From an econometric standpoint, employing factor models to monitor macroeconomic conditions stems from the fundamental insight that information from various aspects and sectors of the economy can be viewed as imperfect measures of a latent common business cycle factor. A robust finding in this area of research is that a few common factors can effectively capture the key features of business cycle fluctuations. This discovery, first documented by Sargent and Sims (1977), has been further supported by using high-dimensional macroeconomic data, as demonstrated in Giannone et al. (2004); Watson (2004).

Following Giannone et al. (2008), DFM based nowcasting framework became the workhorse model of short-term forecasters (Liu et al. (2012), Giannone et al. (2013) Dias et al. (2015), Rusnák (2016), Jiang et al. (2017), Bragoli and Fosten (2018), Caruso (2018)) at many central banks and other institutions. The framework can accommodate a potentially large number of variables by summarizing the information with a few common factors, thus overcoming the so-called curse of dimensionality (Stock and Watson (2002); Bernanke and Boivin (2003)).

Vector autoregression (VAR) models are also widely used in macroeconomics for jointly modelling the dynamics of economic variables. In these general linear models, each variable depends on its own past values and the past values of all other variables, with the pattern of correlation in forecast errors left unconstrained. Bayesian VARs (BVARs) add a layer of complexity by incorporating a naive prior model that assumes all variables are independent white noise or random walks. Early proponents of VAR models in economics, such as Sims (1980); Doan et al. (1984), advocated for the use of Bayesian VARs. Recent research has shown that Bayesian VARs are closely connected with factor models and are well-suited for analysing big data, as demonstrated by De Mol et al. (2008); Banbura et al. (2010b); Antolin-Diaz et al. (2017, 2024).

These methods offer robust statistical frameworks for aligning monthly data with quarterly targets and refining real-time economic assessments and policy decisions. Advancements in these models have enabled the development of automated platforms for real-time monitoring of macroeconomic conditions. Economists delve into a wealth of economic data released by statistical agencies, private and public surveys, and various sources to evaluate the state of the economy. Over time, multiple approaches have developed and utilised to address the significant challenge of distinguishing meaningful signals from noise.

A prominent example of nowcasting using big data is the New York Fed Staff Nowcast, introduced in April 2016 (Aarons et al. (2016)). In the pre-COVID period, this platform issued GDP growth estimates for the current and subsequent quarters each Friday, leveraging up-to-date data releases to refine its predictions. This nowcasting model extracts the latent factors that drive movements in the data and produces a forecast of each economic series that it tracks. If the actual release for that series differs from the model's forecast, this 'news' impacts the current forecast of GDP growth, mirroring the process of market analysts and policymakers, who continually update economic assessments based on new information. However, nowcasting the GDP has become increasingly complex in the aftermath of the COVID-19 pandemic, given the uncertainty and structural changes that have emerged during this period.

The New York Fed Staff Nowcasts 2.0 (Baker et al. (2023)) was introduced in September 2023 to address the shortcomings of the model during the pandemic. To better handle the extreme data releases seen during the pandemic, this paper introduced additional blocks of data (nominal and COVID factors). In addition, the new model introduced stochastic volatility and outlier adjustment to the latent variable dynamics. The nonlinear dynamics in the factor update equations reduce the model's sensitivity to large shocks, ensuring that it can handle smaller surprises as normal without overreacting to the drastic deviations observed during the COVID-19 pandemic. The point forecast was estimated using a Bayesian estimation approach, similar to Antolin-Diaz et al. (2024).

As previously discussed, GDP forecasting in the Indian context has gained attention over the past decade. Notable attempts at nowcasting GDP or Gross Value Added (GVA) in India using the dynamic factor include studies by Bragoli and Fosten (2018); Iyer and Gupta (2019a); Bhadury et al. (2021); Ghosh and Ranjan (2023); Kaustubh et al. (2024); Bhattacharya et al. (2019, 2023). In this paper, we construct a multi-factor dynamic model to estimate India's GDP using soft indicators, nominal factors, and the labour market related indicators, in addition to the real high-frequency economic activity indicators, and the model also provides the source of change in GDP nowcast. However, unlike Fed Staff Nowcasts 2.0 (Baker et al. (2023)), we have adopted a frequentist approach and have normalized the drastic deviations observed during the COVID-19 by either using the stringency index, or using nonlinear data transformation.

3 Literature Gap and Contributions of the study

The literature on nowcasting Indian GDP has primarily focused on single-factor models (Bhadury et al. (2021)), where HFIs related to real economic activity are compressed into a single monthly economic activity factor to calculate the GDP growth nowcast. However, many other types of HFIs related to the Indian economy are available, such as survey-based data (e.g., PMI), cost and supply chain pressure indicators, and labor market indicators, etc. Including these additional HFIs in the nowcasting model can enhance its performance. However, a single-factor approach may not be ideal in this case due to the potential lead-lag relationships among different data types. In this regard, given the advantages of a multi-factor GDP nowcasting model in the presence of various types of high-frequency indicators (HFIs), Bhattacharya et al. (2019, 2023) developed multi-factor models for nowcasting Indian GDP. They used PCA to create multiple factors by compressing the maximum variations of HFIs into a few factors. However, unlike Bhattacharya et al. (2019, 2023), our approach involves creating multiple factors based on economic intuition rather than relying solely on statistical techniques like PCA. We focus on factors such as real economic activity, cost pressures, survey-based data, and labor market data. This approach allows for richer macroeconomic surveillance, as in addition to calculating GDP growth nowcast, we can also track monthly movements in these factors to provide a summarized view of real economic activity, cost pressures, survey-based data, and the labour market. In contrast, factors derived mechanically from statistical techniques like PCA are often difficult to interpret economically and primarily serve to provide final GDP nowcast estimates. The above-explained factors are modeled in a mixed-frequency multi-factor vector autoregressive (VAR) framework to capture the lead-lag relationship between the factors.

The other downside of the earlier nowcasting models for the Indian economy is their singular focus on the final aggregate GDP nowcasts. However, as policymakers increasingly prioritize a more detailed analysis following revisions of nowcasts after the release of new HFI data, there is a growing interest in understanding the precise sectors or HFIs responsible for these revisions. As far as we know, there hasn't been any GDP nowcasting model for the Indian economy in existing literature that allows for a breakdown of the changes in GDP nowcasts following a new data release. This breakdown is necessary to pinpoint and measure the influence of data revisions and new data releases of HFIs on GDP nowcasts. In this paper, we aim to quantify the impact of any new data release or revision in the pre-released data and decompose the effects of data releases on the nowcasts for each HFI. This approach allows us to precisely discern which surprises in specific HFIs are driving the revisions in the nowcast.

In our analysis, we also introduce additional indicators to capture the divergence between predicted growth based on real economic activity and actual GDP growth affected by price fluctuations and supply chain disruptions. Indicators such as the ratio of CPI to WPI, the Baltic dry index, and the IMF commodity price index are employed to gather information related to supply shocks impacting GDP.

The New York Fed Staff Nowcasts 2.0 (Baker et al. (2023)) deals with outliers in the HFIs momentum due to economic disruptions like COVID through non-linear Bayesian DFM estimation incorporating stochastic volatility and outlier detection in the model. However, in our model, we have transformed data using sigmoid function before the DFM estimation to ameliorate the impact of outliers in HFIs momentum in the estimation. Additionally, in the model with no transformation, we utilize the Oxford Stringency Index to account for disruptions caused by lockdowns or movement restrictions, reducing the impact of outlier data during the pandemic on parameter estimation. Both approaches have effectively minimized the impact of outliers, leading to improved nowcasts in the post-pandemic period.

To summarize, our contributions are twofold. Firstly, in the literature on nowcasting Indian GDP, we decompose the GDP nowcast using blocks of real, nominal, survey, and labor data, and examine the impact of new data releases on GDP. Secondly, on the methodological front, we use a sigmoid transformation to reduce the model parameters' sensitivity to large shocks.

4 Methodology

4.1 Data Transformation

Major disruptions like COVID-19 interfere with the data-generating process, causing outliers in both the target variable and HFIs[3]. These disruptions make the parameters highly unstable and sensitive to shocks. To reduce this sensitivity to large shocks, we make the following transformation. We first standardized the momentum of HFIs[4] and then apply the sigmoid transformation on the standardized momentum using equation 1 and 2. Equation 1 is a linear transformation (also called *z-tranformation*) of the Y_i, the momentum of i^{th} HFI, and is computed by subtracting its mean $\mu(Y_i)$ and then dividing it by its standard deviation $\sigma(Y_i)$. Equation 2 is nonlinear transformation applied on the *z-tranform* of the HFI momentum. As seen in equation 2, we have introduced a non-negative parameter w_i that is first multiplied with the *z-tranform* of the HFI momentum of i^{th} variable before applying Sigmoid transformation. The parameter w_i determines the transformation's sensitivity to outliers, with a higher w_i data becomes less sensitive to outliers (Figure 1).

$$z(Y_i) = \frac{Y_i - \mu(Y_i)}{\sigma(Y_i)} \tag{1}$$

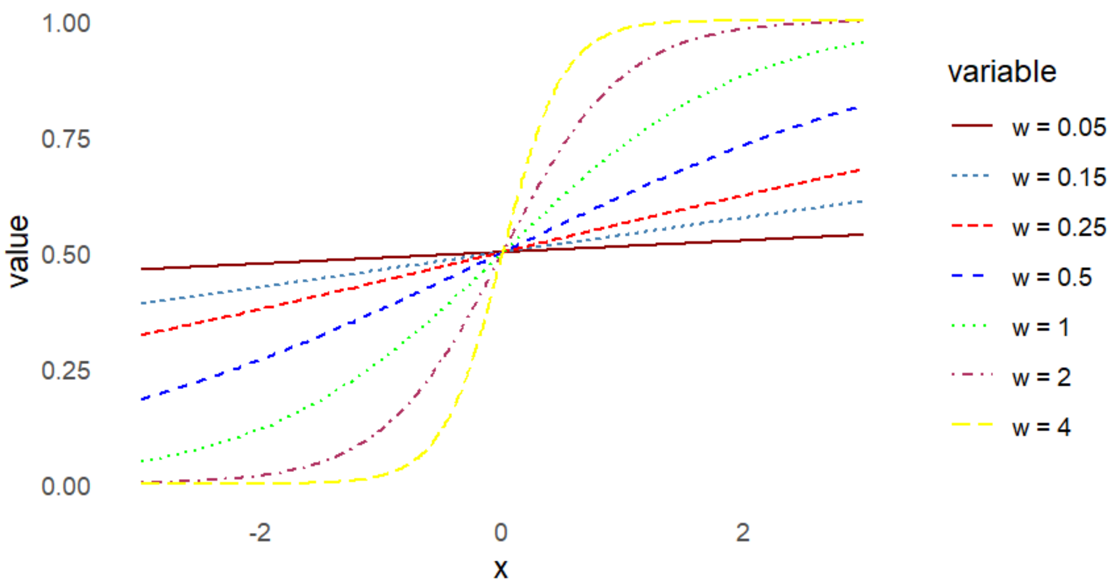
[3]Data source and their availability period is described in Table 9 of the Appendix

[4]All the HFIs are deseasonalized before computing the momentum

$$y_i = Sigmoid(w_i * z(Y_i)) = \frac{1}{1 + exp(-w_i * z(Y_i))} \text{ where } w_i > 0 \tag{2}$$

As discussed, to mitigate the impact of outlier events like COVID-19 on the model's overall forecasting performance, we transform the momentum of the HFIs using Sigmoid function. The smaller value of the transformation parameter w_i implies impact of outliers is high in the model, whereas higher w_i decreases the impact of the outlier data. Sigmoid function is a function from $(-\infty, \infty)$ to $(0, 1)$. For larger value of x, it suppresses the data close to 1. This helps in suppressing outlier value, while maintaining properties of linear mapping for values close to $x = 0$.

Figure 1: Sigmoid Transformation at different w



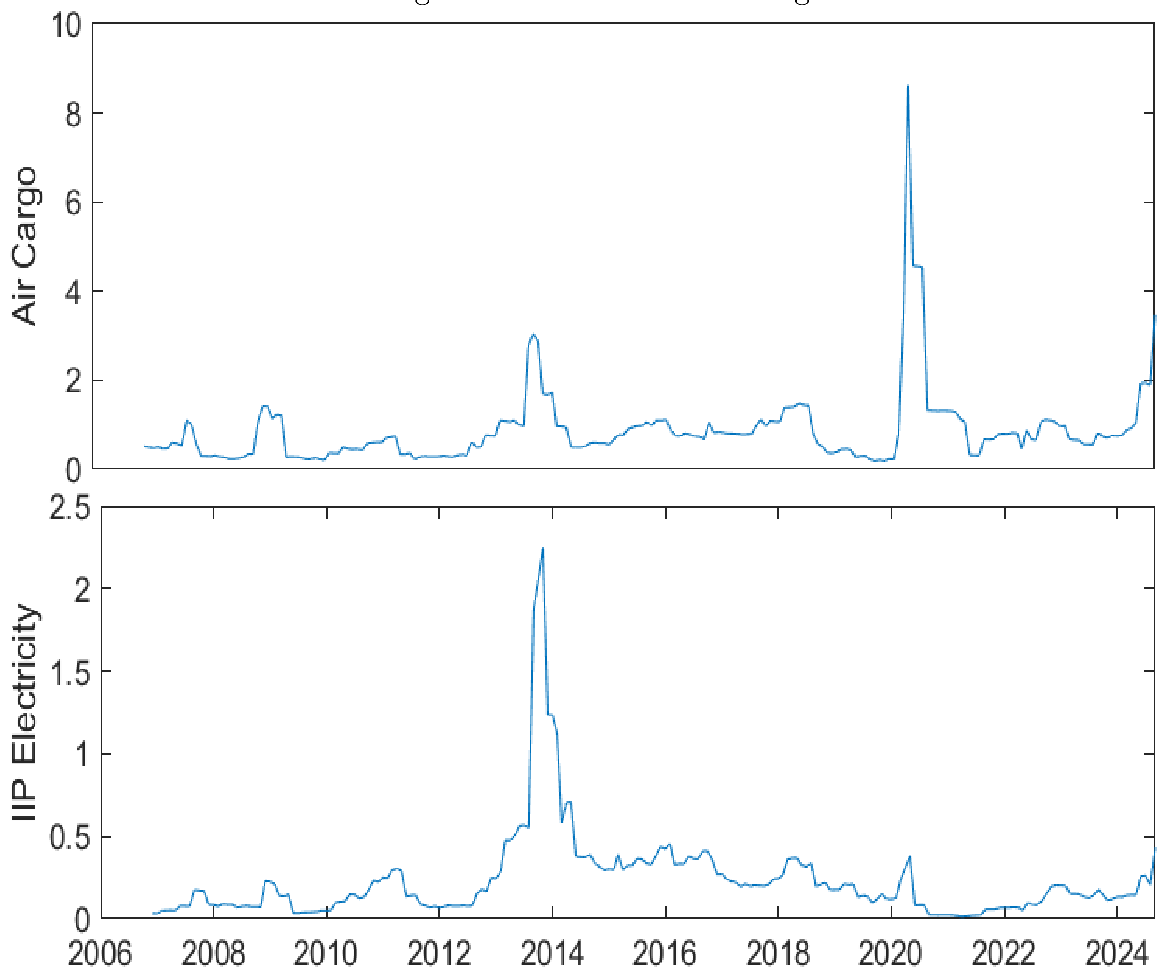
To find the value of the parameter w_i, we have used analytical approach, in which we have chosen the value of w_i based on the ratio of volatilty between the momentum of HFI Y_i and the momentum of GDP in both long run and in the recent past (See Equation 3). The ratio is chosen such that the higher volatile series flatten out, and the highly volatile HFI has lower contribution to the GDP growth nowcast. While transforming the momentum, we consider both series variance as well as short term variance. It reduces the impact of HFIs which has high volatility in the last 12 months. This helps in ameliorating the impact of one time shocks such as COVID-19 on the overall forecasting performance of the model.

$$w_i = \frac{\sigma(Y_i)\sigma'(Y_i)}{\sigma(GDP)\sigma'(GDP)} \tag{3}$$

where $\sigma(Y_i)$ and $\sigma(GDP)$[5] are the standard deviation of Y_i and GDP, respectively, and, $\sigma'(Y_i)$ and $\sigma'(GDP)$ are the standard deviation of Y_i and GDP, respectively, over last one year.

Figure 2 illustrates the transformation weight for Air Cargo and IIP Electricity at different time. During 2013-14, there was a simultaneous shock to Air Cargo and IIP Electricity, and the transformation to the series is similar for both HFIs. However, during Covid Shock, there was forced lockdown contributing to very high volatility in Air Cargo, and its impact was reduced for nowcasting GDP, however the volatility of IIP electricity didn't see such drastic change.

Figure 2: Transformation Weight



[5]Our primary focus is on improving the GDP nowcast. To improve the GVA nowcast, one should use the standard deviation of GVA.

4.2 State Space Description

We nowcast GDP using a mixed frequency (monthly and quarterly) multi-factor in a state space framework using an expectation maximization algorithm similar to Banbura et al. (2010a) for both trnasformed data and with an additional HFI (Oxford stringency index) for the untransformed data .

The state space representation of the model is as follows: The $y_{i,t}$ denotes month-on-month (M-o-M) percentage change of the i^{th} monthly high frequency indicator at time t. $f_{1,t}, \cdots, f_{r,t}$ are the unobserved factors at time t that are linked to the HFIs through the observation equations shown in equation 4. $\varepsilon_{i,t}$ is the idiosyncratic error for the i^{th} HFI and $\gamma_{i,r}$ denotes the factor loading of the r^{th} factor on the i^{th} HFI.

$$y_{i,t} = \mu_i + \gamma_{i,1} * f_{1,t} + \cdots + \gamma_{i,r} * f_{r,t} + \varepsilon_{i,t} \tag{4}$$

The unobserved factors are modelled as VAR with order p[6], as shown in the transition equation below (equation 5), where f_t denotes the vector of all factors and $a_1, a_2, \cdots, a_p$ are the matrices of the autoregressive components of the VAR.

$$f_t = a_1 * f_{t-1} + \cdots + a_p * f_{t-p} + \eta_t, \;\; \eta_t \sim N(0, \sigma_f^2) \tag{5}$$

Finally, the idiosyncratic errors of the HFIs $\varepsilon_{i,t}$ in equation 4 are modelled as an $AR(1)$ process, as shown in equation 6 below.

$$\varepsilon_{i,t} = \rho_i * \varepsilon_{i,t-1} + \xi_{i,t}, \;\; \xi_{i,t} \sim N(0, \sigma_i^2) \tag{6}$$

We have modelled quarterly variables similar to Banbura et al. (2010a). The quarterly variable, like GDP, at the quarterly level denoted as GDP_t^Q, $t = 3, 6, 9, \cdots$ is modeled as the sum of unobserved monthly contributions GDP_t^M, as shown in equation 7.

$$GDP_t^Q = GDP_t^M + GDP_{t-1}^M + GDP_{t-2}^M \tag{7}$$

Following Mariano and Murasawa (2003), the QoQ percentage change of GDP_t^Q, y_t^Q can be approximated to a linear combination of lags of MoM momentum of unobserved monthly contributions GDP_t^M, y_t as described in equation 8.

$$y_t^Q = y_t + 2 * y_{t-1} + 3 * y_{t-2} + 2 * y_{t-3} + y_{t-4} \tag{8}$$

[6]In our model, $p = 4$

The momentum of unobserved monthly contributions GDP_t^M, y_t is modelled to have the same factor representation as the monthly HFIs shown in equation 9 and 10.

$$y_t = \mu_Q + \gamma_1 * f_{1,t} + \cdots + \gamma_r * f_{r,t} + \varepsilon_t^Q \tag{9}$$

$$\varepsilon_t^Q = \rho_Q * \varepsilon_{t-1}^Q + \xi_t^Q,\ \ \xi_t^Q \sim N(0, \sigma_Q^2) \tag{10}$$

The model is estimated using the Expectation Maximization (EM) algorithm using Maximum Likelihood, as in Banbura et al. (2010a). The EM algorithm begins by calculating the principal components of the HFIs listed in the given factor, and the model parameters are then estimated using OLS regression, treating the principal components as if they were the true common factors. As the principal components are reliable estimates of common factors for large datasets, it is the initialization point for the algorithm. In the next step, using the estimated parameters, an updated estimate of the common factors and model parameters is obtained through the Kalman smoother. The process stops when convergence is reached by maximizing the log-likelihood.

4.3 Factors and Indicators description

The nowcasting model in the study has four unobserved factors. The first factor, the real economic activity index, loads the monthly HFIs that represent real economic activity. They are primarily volume indicators of real economic activity. The nominal index, the second factor, loads HFIs with information about prices and global supply chain pressures. The third factor, the Soft Index, loads survey-based data, such as PMI Manufacturing and Services. Finally, we load HFIs related to the labour market into the fourth index, the labour index. The selection of factors and HFIs listed are similar to that of Baker et al. (2023). However, unlike Baker et al. (2023), who load all HFIs into a global factor, we load only HFIs related to real economic activity into our factor, leaving out high-frequency indicators related to prices, surveys, and the labour market. This allows us to model the lead-lag relationships between HFIs in the real economy, the labour market, the nominal economy, and survey-based HFIs without overlapping in the VAR setting.

Table 1 provides a comprehensive list of HFIs and quarterly indicators used in the model. The table provide details on the transformation applied to each indicator prior to their inclusion in the model, as well as the list of indicators loaded into each factor.

Table 1: HFIs and Quarterly Indicators

Series Name	Frequency	Real Activity	Nominal	Survey	Labour	Transformations
Domestic Air Cargo	Monthly	✓				Monthly % Change
Auto Sales	Monthly	✓				Monthly % Change
IIP Consumption	Monthly	✓				Monthly % Change
Nong Import	Monthly	✓				Monthly % Change
Rail Freight	Monthly	✓				Monthly % Change
Goods Export	Monthly	✓				Monthly % Change
Foreign Tourists	Monthly	✓				Monthly % Change
IIP Core	Monthly	✓				Monthly % Change
Domestic Air Passengers	Monthly	✓				Monthly % Change
IIP Cement	Monthly	✓				Monthly % Change
IIP Electricity	Monthly	✓				Monthly % Change
IIP Manufacturing	Monthly	✓				Monthly % Change
Rail Passengers	Monthly	✓				Monthly % Change
PMI Manufacturing	Monthly			✓		Level
PMI Services	Monthly			✓		Level
Crude Steel Consumption	Monthly	✓				Monthly % Change
Petrol Consumption	Monthly	✓				Monthly % Change
Vehicle Registration	Monthly	✓				Monthly % Change
GST Revenue	Monthly	✓				Monthly % Change

Continued on next page

Table 1: HFIs and Quarterly Indicators (Continued)

Series Name	Frequency	Real Activity	Nominal	Survey	Labour	Transformations
Sea Cargo	Monthly	✓				Monthly % Change
IIP Capital Goods	Monthly	✓				Monthly % Change
Tractor Sales	Monthly	✓				Monthly % Change
Electricity Demand	Monthly	✓				Monthly % Change
Stringency Index	Monthly	✓				Level Change
Ratio CPI to WPI	Monthly		✓			Monthly % Change
Baltic Dry Index	Monthly		✓			Monthly % Change
IMF Commodity Price Index	Monthly		✓			Monthly % Change
Naukri Jobspeak	Monthly				✓	Monthly % Change
CMIE Urban Unemployment	Monthly				✓	Monthly % Change
New EPFO Subscribers	Monthly				✓	Monthly % Change
CMIE Rural Unemployment	Monthly				✓	Monthly % Change
E-Way Bills	Monthly	✓				Monthly % Change
Hotel Occupancy Rate	Monthly	✓				Monthly % Change

Continued on next page

Table 1: HFIs and Quarterly Indicators (Continued)

Series Name	Frequency	Real Activity	Nominal	Survey	Labour	Transformations
Average Revenue Per Available Room	Monthly	✓				Monthly % Change
Gross Domestic Product	Quarterly	✓	✓	✓	✓	Quarterly % Change
Gross Value Added	Quarterly	✓	✓	✓	✓	Quarterly % Change

4.4 Nowcast Evaluation using Root Mean Square Error (RMSE)

The nowcast evaluation is based on an out-sample one-period ahead rolling window evaluation using the root mean square error (RMSE) as the metric for the time period between *2021: Q1 and 2024: Q2*. We have evaluated the model performance when we have 1 month of HFIs data, 2 months of HFIs data, and all 3 months of HFIs data for the quarter for which evaluation is being conducted. To highlight the importance of incoming HFIs data in improving the nowcasting performance, we have also calculated the RMSE when there is no data on HFIs for the given quarter and the nowcast is based on HFIs data only up to the previous quarter. The RMSE is computed for the Benchmark Model ARIMA, single-factor and Multi-Factor Models for both untransformed and transformed data. The robustness of the model comparision is further tested using Diebold and Mariano Test (Diebold and Mariano (1995); Diebold (2015)).

4.5 News and Nowcast update

As shown in equation 11, the nowcast revision is expressed as a weighted sum of news from the released data of various indicators. The nowcast revision is given on the left-hand side of equation 11, where Ω_ν denotes the information available at time ν and $\Omega_{\nu+1}$ denotes the information available up to time $\nu+1$. $\Omega_\nu \subseteq \Omega_{\nu+1}$ and $\Omega_{\nu+1} \setminus \Omega_\nu$ is the set of news released between time ν and $\nu+1$. Assuming $\nu+1$ is the next release date for indicators following time ν, then $\Omega_{\nu+1} \setminus \Omega_\nu$ comprises the news from the set of indicators released on date $\nu+1$. $\mathcal{J}_\nu$ denotes the set of indicators with release date ν.

Let $\mathbb{E}[y_t^Q|\Omega_{\nu+1}]$ denotes the prediction about y_t^Q given the information set $\Omega_{\nu+1}$ and $\mathbb{E}[y_t^Q|\Omega_\nu]$ denotes the prediction about y_t^Q given the information set Ω_ν. The difference between them is the change in the nowcast from the time ν to $\nu+1$, or the impact of news from the indicators $\mathcal{J}_{\nu+1}$ released on date $\nu+1$. The news is defined as the change in the expected value of the j^{th} HFI for date t between time ν and $\nu+1$. Similarly $x^j_{t,\nu} = \mathbb{E}[x^j_t \mid \Omega_\nu]$ is the projected value of the j^{th} HFI at time t, as of time ν and x^j_t is the true value released on date $\nu+1$. The weights $b_{j,t,\nu+1}$[7] given to each indicator is determined using the Kalman gain which depends on signal to noise ratio of the HFI associated with each indicator (Banbura et al. (2010a)).

$$\mathbb{E}[y^Q_t|\Omega_{\nu+1}] - \mathbb{E}[y^Q_t|\Omega_\nu] = \sum_{j\in J_{\nu+1}} b_{j,t,\nu+1}(x^j_{t,\nu+1} - \mathbb{E}[x^j_t \mid \Omega_\nu]) \tag{11}$$

The right-hand side of equation 11 denotes the weighted sum of the difference between released value and expected value of the HFIs released on date $\nu+1$. We use equation 11 to compute the impact of data releases on GDP nowcast at different time.

5 Results

5.1 DFM Results

Table 2 shows the factor loadings of the monthly HFIs for all four factors explained earlier. The factor loadings for the Real Economic Activity Index largely align with expectations, showing negative values for the Stringency Index, and positive values for the HFIs that reflect real economic activity. As previously discussed, the nominal index reflects the pressures on the supply chain in the economy, resulting in positive loadings for the Baltic dry index and commodity prices, and a negative loading for the ratio of CPI to WPI. This divergence between CPI and WPI in India, we hypothesize, indicates higher profitability due to lower input costs and higher output prices. As expected, the soft indicators have positive weights for PMI indicators both manufacturing and services. Finally, the Labour Index has positive weights for Naukri Jobspeak job postings and negative weights for rural and urban unemployment numbers. A positive factor loading of the monthly HFI indicates that, as the momentum of the HFI increase, the GDP momentum is also expected to increase, if the factor loading of that factor for GDP is positive and vice versa if the factor loading of that factor for GDP is negative. Similarly, a negative loading of the monthly HFI means that when the momentum of HFI increases, the GDP momentum is expected to fall, if the factor loading of that factor for GDP is positive and vice versa

[7]It is worth noting that for t large enough so that the Kalman filter has approached its steady state, the weights $b_{j,t,\nu+1}$ will be uniquely defined and doesn't depend on a particular realisation of x^j under the Gaussian assumption for x^j. Please refer to equation 13 of Antolin-Diaz et al. (2024) for the computation of $b_{j,t,\nu+1}$. The Gaussian assumption is violated when we transform the data using sigmoid transformation, and therefore the impact of each HFIs can't be computed using the transformed data in case of simultaneous release of data.

if the factor loading of that factor for GDP is negative. An incorrect sign can lead to a misinterpretation of the relationship between the target variable GDP and the latent factor HFI.

Table 2: Factor Loadings for Monthly Series

Activity Index Type	Series Name	Loading (Untransformed)	Loading (Transformed)
Real	Domestic Air Cargo	0.24	0.25
	Auto Sales	0.25	0.11
	IIP Consumption	0.28	0.30
	Nong Import	0.12	0.14
	Rail Freight	0.22	0.28
	Goods Export	0.21	0.16
	Foreign Tourists	0.17	0.19
	IIP Core	0.25	0.32
	Domestic Air Passengers	0.24	0.17
	IIP Cement	0.26	0.28
	IIP Electricity	0.18	0.29
	IIP Manufacturing	0.29	0.30
	Crude Steel Consumption	0.21	0.12
	Petrol Consumption	0.24	0.26
	Vehicle Registration	0.09	0.06
	GST Revenue	0.16	0.11
	Sea Cargo	0.07	0.14
	IIP capital Goods	0.25	0.18
	Tractor Sales	0.23	0.09
	Electricity Demand	0.18	0.27
	Stringency Index	-0.17	NA
	E-Way Bills	0.16	0.16
	Hotel Occupancy Rate	0.11	0.14
	Average Revenue Per Available Room	0.08	0.13

Continued on next page

Table 2: Factor Loadings for Monthly Series (Continued)

Activity Index Type	Series Name	Loading (Untransformed)	Loading (Transformed)
Nominal	Ratio CPI to WPI	−0.63	−0.69
	Baltic Dry Index	0.41	0.27
	IMF Commodity Price Index	0.65	0.67
Soft	PMI Manufacturing	0.71	0.71
	PMI Services	0.71	0.71
Labour	Naukri Jobspeak	0.60	0.86
	CMIE Urban Unemployment	−0.49	−0.29
	New EPFO Subscribers	0.37	0.34
	CMIE Rural Unemployment	−0.50	−0.22

Table 3 shows the factor loadings of each factor for the quarterly indicators. The signs of the loadings are according to expectation, as the loading for the Real Economic Activity Index is positive. The loading for the nominal index is negative as it indicates the cost pressures in the economy. The loadings for the soft index and labour index are positive, as expected. Note that the quarterly indicators load on four lags of the indicators, unlike the monthly HFIs, which only load on contemporaneous factors. However, the loading of the lag of the factors are the multiples of the loadings of the contemporaneous factors based on equation 8.

Table 3: Factor Loadings for Quarterly Variables (Untransformed)

Quarterly Indicators	Real Activity Factor	Nominal Factor	Soft Factor	Labour Factor
Gross Domestic Product	0.05	−0.02	0.01	0.12
Gross Value Added	0.05	−0.02	0.01	0.12

Table 4: Factor Loadings for Quarterly Variables (Transformed)

Quarterly Indicators	Real Activity Factor	Nominal Factor	Soft Factor	Labour Factor
Gross Domestic Product	0.08	−0.02	0.01	0.03
Gross Value Added	0.08	−0.02	0.01	0.03

Figure 3 shows the movement of factors over time for untransformed data[8]. The top left chart shows the movement of the Real Economic Activity Index in the dark color along with the movement of other HFIs in colored shades. The movement of the Real Activity Index captures the movement of overall economic conditions well. It captures the dip in economic conditions during the global financial crisis. It also depicts the negative economic momentum resulting from the lockdown during the COVID pandemic, followed by recovery upon resumption of normal operations. The top-right chart shows the nominal index's movement over time. It captures well the increase in cost pressures during the start of the Ukraine war and the easing of those pressures in 2023. The chart on the left below illustrates the movement of the Soft Index, which closely mirrors the movement of the PMI indices. It exhibits a significant decline during the COVID-19 pandemic and then resumes its expansionary trend during the recovery period. The below right chart shows the movement of the Labour Index with time.

[8]For transformed data, please see Figure 6 in Appendix

Figure 3: Factors over Time (Untransformed)

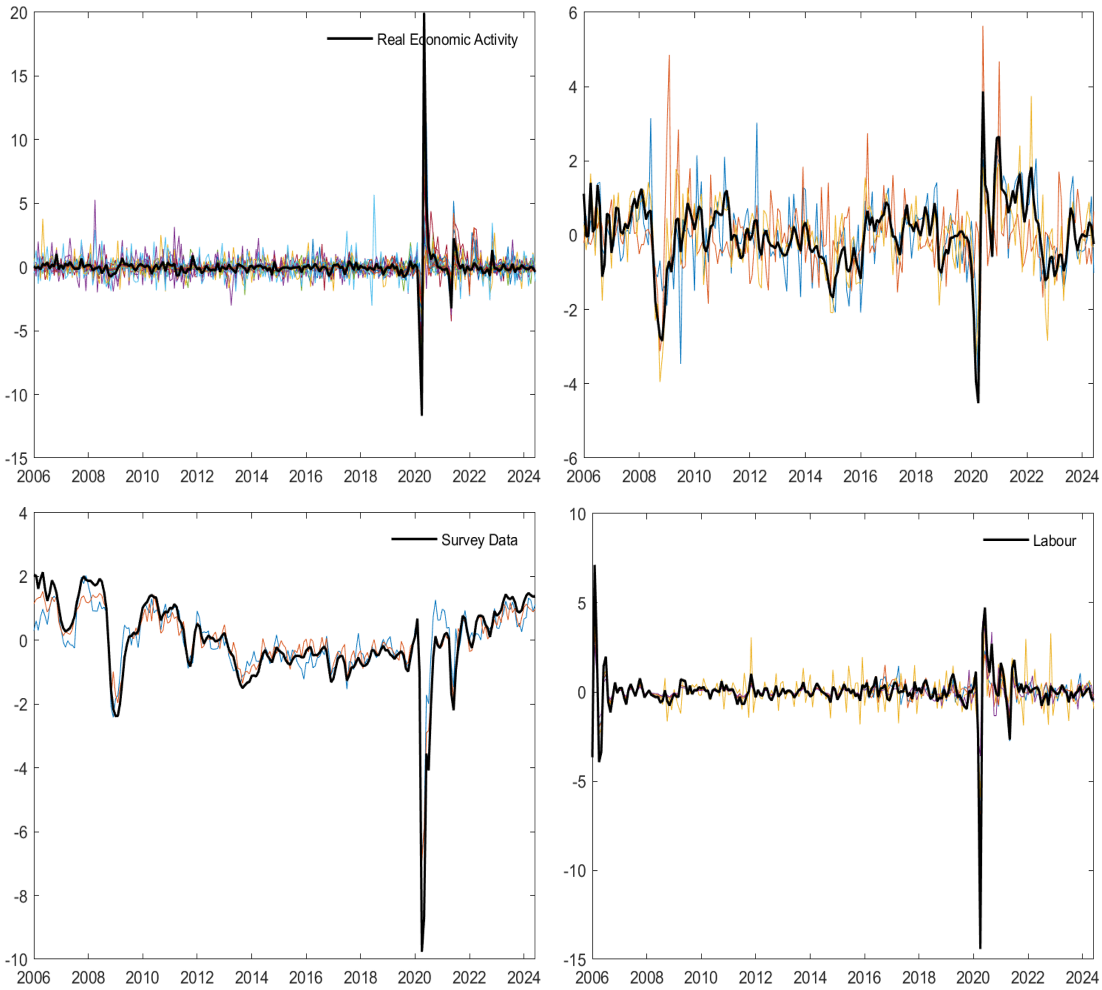
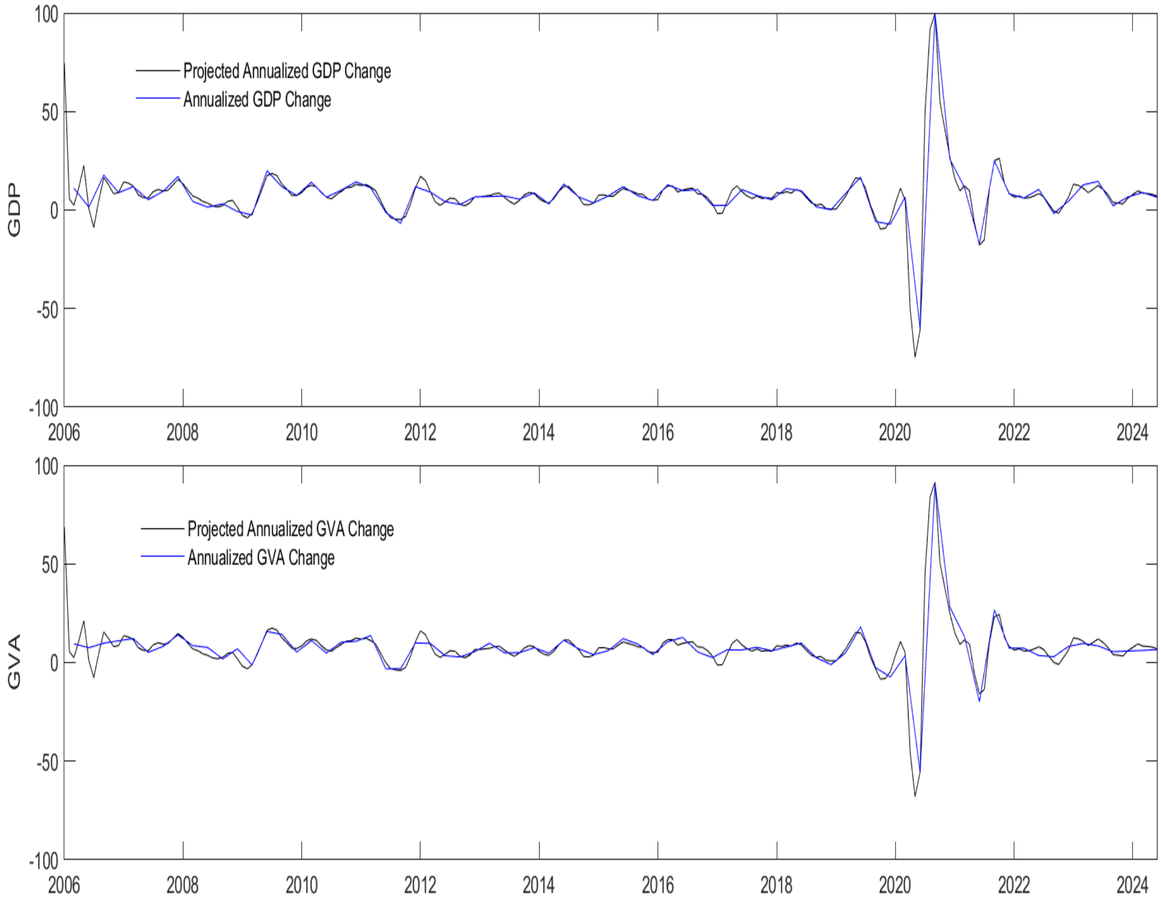


Figure 4 shows the movement of GVA and GDP QoQ momentum along with their projected QoQ momentum based on estimated factors (the sum product of the factors and their respective factor loadings for GDP and GVA). Figure 4 shows that factors track the momentum of GDP and GVA well for untransformed data[9].

[9]For transformed data, please see Figure 7 in Appendix

Figure 4: Comparison of movement of QoQ momentum of GVA and GDP with their projected QoQ momentum, based on factors (Untransformed)



5.2 Out-of-Sample Nowcast Evaluation Results

The results of the out-of-sample one-period ahead rolling window evaluation of the quarter-on-quarter (QoQ) annualized growth nowcasts of GDP and GVA in the post-pandemic period can be found in Table 5. The test sample for our model consists of 14 quarters, from *2021: Q1 to 2024: Q2*. The test sample period contains second wave of Covid. In both scenarios, RMSE decreases substantially for GDP with the inclusion of more HFI data points for the quarter for which the growth is nowcasted. This shows that the model has been able to include the incoming information of HFIs well in nowcasting GDP growth. Table 6 compares a benchmark model ARIMA, a single-factor model with only HFIs related to the real economy, a single-factor model with all HFIs, and then more substantially for the multi-factor model with all HFIs for both untransformed and transformed data. Compared to ARIMA, the QoQ annualized RMSE decreases sufficiently for the single-factor models with only real activity HFIs. This further decreases for the single-factor model with all HFIs and multi-factor model with all HFIs. This improves further when data is transformed, justifying the use of the multi-factor model with transformed data for nowcasting GDP.

Table 5: Out-of-sample RMSE QoQ (Annualized)

	Untransformed		Transformed	
	GDP	GVA	GDP	GVA
0 Month Data	13.8	14.4	9.5	9.6
1 Month Data	8.1	9.0	6.4	8.1
2 Month Data	3.8	4.1	4.2	4.6
3 Month Data	3.5	4.7	2.9	4.4

Table 6: Out-of-sample Evaluation of GDP nowcast for different models

Model	Untransformed	Transformed
ARIMA	40.0	40.0
Single-Factor Model with Real Activity HFIs	4.5	4.0
Single-Factor Model with all HFIs	4.4	3.9
Multi-Factor Model with all HFIs	3.5	2.9

We further test for robustness of the above findings. We use Diebold and Mariano test (Diebold and Mariano (1995); Diebold (2015) to test that the difference in the RMSE is significant, and our findings are robust. Our test hypothesis is:

Null Hypothesis H_0 : Both Models has equal forecast accuracy.

Alternate Hypothesis H_1 : Model 2 has better forecast accuracy than Model 1.

The test results are summarized in Table 7.

Table 7: P-Value for different robustness test

Model 1	Model 2	p-value (DM)
ARIMA	Multi-Factor Model All HFIs (Untransformed)	0.14

Continued on next page

Table 7: P-Value for different robustness test (Continued)

Model 1	Model 2	p-value (DM)
ARIMA	Multi-Factor Model All HFIs (Transformed)	0.14
Single-Factor Model Real Activity (Untransformed)	Multi-Factor Model All HFIs (Untransformed)	0.15
Single-Factor Model All HFIs (Untransformed)	Multi-Factor Model All HFIs (Untransformed)	0.16
Single-Factor Model Real Activity (Transformed)	Multi-Factor Model All HFIs (Transformed)	0.19
Single-Factor Model All HFIs (Transformed)	Multi-Factor Model All HFIs (Transformed)	0.18
Single-Factor Model Real Activity (Untransformed)	Single-Factor Model Real Activity (Transformed)	0.18
Single-Factor Model All HFIs (Untransformed)	Single-Factor Model All HFIs (Transformed)	0.15
Multi-Factor Model All HFIs (Untransformed)	Multi-Factor Model All HFIs (Transformed)	0.37

The two important findings of the test are i) Multi-Factor Model improves the nowcast for GDP significantly, and ii) Nowcast with transformed data are better across different models.

5.3 Nowcast News and Updates

Figure 5 shows the nowcast of *Q2: 2024* YoY growth for various data vintages for GDP using multi-factor model with both transformed and untransformed data. Each HFI updates the nowcast, and the model allows us to pinpoint the specific HFI that influences the GDP nowcast with new data releases. Furthermore, Table 8 shows the decomposition of change in the QoQ GDP nowcast for model with untransformed data[10] for *Q2: 2024* from the data vintage of *May 15, 2024* to the data vintage of *June 25, 2024*, *June 25, 2024* to *July 16, 2024*, and *July 16, 2024* to *August 28, 2024* due to new data releases from different HFIs. This table lists the impact of each HFI on the QoQ GDP nowcast for Q2 2024 following their data release. We list the impact of each indicator in the column during the respective timeframe. The impact is calculated by multiplying the surprises

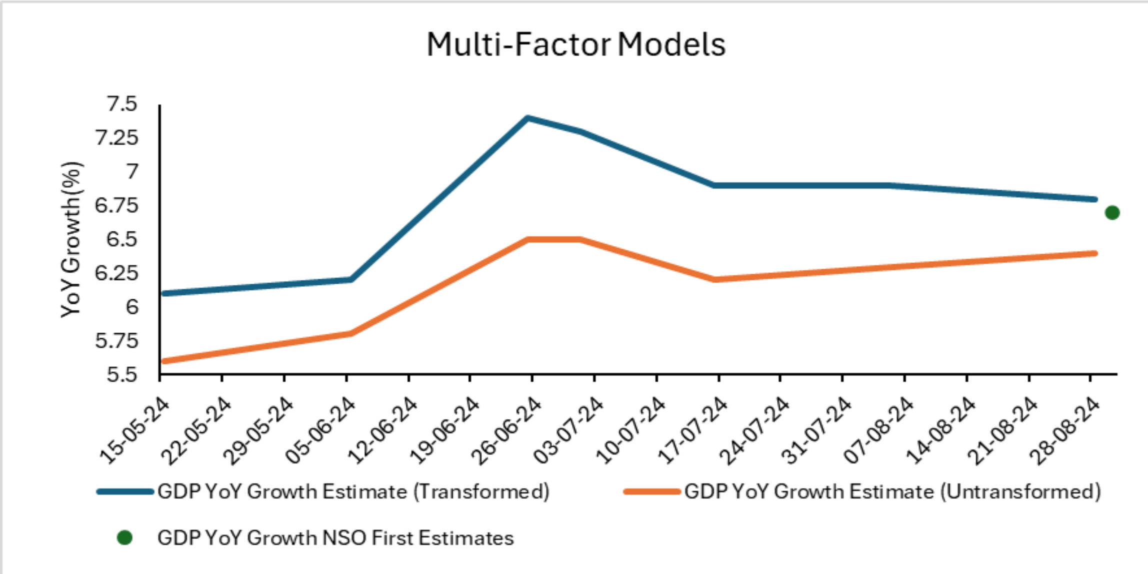
[10]This decomposition is not possible for transformed data

in the indicators by their respective weights.

Table 8: Impact of HFIs on change in GDP Nowcasts between different time periods

HFIs	15 May to 25 June	25 June to 16 July	16 July to 28 Aug
Domestic Air Cargo	-0.03	0.00	0.00
Auto Sales	-0.11	0.00	0.12
IIP Consumption	0.32	0.12	-0.01
NONG Imports	0.01	0.00	-0.03
Rail Freights	0.03	0.00	0.00
Exports	0.13	-0.02	0.10
Foreign Tourists Arrivals	0.00	0.00	0.00
IIP Core	0.30	-0.01	0.00
Domestic Air Passengers	0.78	-0.37	-0.05
IIP Cement	-0.12	0.00	-0.01
IIP Electricity	0.13	0.04	0.00
IIP Manufacturing	0.57	0.06	0.05
PMI Manufacturing	1.57	0.00	0.04
PMI Services	-0.17	-0.17	0.24
Crude Steel Consumptio	0.04	0.00	-0.02
Petrol Consumption	-0.09	-0.01	-0.08
Vehicle Registration	0.01	0.00	0.02
GST Collection	0.66	0.00	0.06
Sea Cargo	0.04	0.00	-0.01
IIP Capital Goods	0.05	0.00	0.00
Tractor Sales	-0.04	0.00	0.02
Electricity Demand	0.18	0.00	0.04
Ratio of CPI to WPI	-0.22	-0.05	0.16
Baltic Dry Index	-0.01	0.00	0.01
IMF Commodity Prices	-0.05	-0.02	-0.01
Naukri Jobspeak	0.62	-0.36	0.00
CMIE Urban Unemployme	-0.23	-0.37	0.01
EPF Subscribers	-0.04	0.00	0.04
CMIE Rural Unemployme	-0.41	-0.26	-0.01
E-WAY Bills	0.14	0.01	0.03
Hotel Occupancy Rate	0.00	0.00	0.00
Revenue Per Room	0.00	0.00	-0.01

Figure 5: Nowcasts of 2024: Q2 GDP at different vintages for multi-factor models



6 Concluding remarks

Many central banks widely employ nowcasting models, which have become increasingly vital tools for addressing uncertainties about the current economic state. Policymakers want accurate forecasts, as well as an understanding of the events that led to the nowcast and any surprises. While there is extensive literature on the decomposition of surprises in nowcasts for western economies, such studies are notably lacking for India. Our study aims to fill this gap by providing insights that assist policymakers not only in measuring point forecasts but also in identifying surprise sources.

In our analysis, we compared dynamic factor models with a single factor to those with multiple factors, where each factor represents a different economic block. Based on our findings, we recommend using a mixed-frequency multi-factor model for nowcasting Indian GDP, as it addresses many challenges posed by nowcasting in the post-pandemic era. This model offers flexibility to incorporate new high-frequency indicators (HFIs) or additional factors as they become available and necessary. Furthermore, it provides richer information by summarizing monthly movements in cost pressures, survey-based data, and labor market conditions, alongside broader economic trends. This approach also allows policymakers to conduct more granular analyses, identifying the factors behind changes in GDP nowcast as new data emerges.

Additionally, we propose an alternative methodology in form of sigmoid transformation to reduce the impact of outliers or shocks in the data-generating process. This improves the overall nowcasting performance for the GDP data, however, it is not possible to attribute the contribution of each HFIs on the GDP nowcast in case of simultaneous data release after sigmoid transformation. A natural extension of this research is to disaggregate the contribution of each HFIs into independent and correlated contribution, in case of simultaneous release of HFIs on GDP. Although our empirical work focuses on India, the framework we have developed can be easily adapted for use in other countries.

References

G. Aarons, D. Caratelli, M. Cocci, D. Giannone, A. Sbordone, and A. Tambalotti. Just released: Introducing the FRBNY nowcast. *Liberty Street Economics*, 2016.

J. Antolin-Diaz, T. Drechsel, and I. Petrella. Tracking the slowdown in long-run GDP growth. *Review of Economics and Statistics*, 99(2):343–356, 2017.

J. Antolin-Diaz, T. Drechsel, and I. Petrella. Advances in nowcasting economic activity: The role of heterogeneous dynamics and fat tails. *Journal of Econometrics*, 238(2): 105634, 2024.

K. Baker, M. Almuzara, H. O'Keeffe, and A. M. Sbordone. Reintroducing the New York Fed Staff nowcast. *Federal Reserve Bank of New York*, 20230908, 2023.

M. Banbura, D. Giannone, and L. Reichlin. Nowcasting. *ECB Working Paper (No. 1275)*, 2010a.

M. Banbura, D. Giannone, and L. Reichlin. Large bayesian vector autoregressions. *Journal of Applied Econometrics*, 25(1):71–92, 2010b.

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

S. S. Bhadury, S. Ghosh, and P. Kumar. Constructing a coincident economic indicator for India: How Well Does It Track Gross Domestic Product? *Asian Development Review*, 38(2):237–277, 2021.

R. Bhattacharya, R. Pandey, and G. Veronese. Tracking India growth in real time. *National Institute of Public Finance and Policy*, page 11/90, 2011.

R. Bhattacharya, P. Chakravartti, and S. Mundle. Forecasting India's economic growth: a time-varying parameter regression approach. *Macroeconomics and finance in emerging market economies*, 12(3):205–228, 2019.

R. Bhattacharya, B. Bhandari, and S. Mundle. Nowcasting India's quarterly GDP growth: A Factor-Augmented Time-Varying Coefficient Regression Model (FA-TVCRM). *Journal of Quantitative Economics*, 21(1):213–234, 2023.

D. Bragoli and J. Fosten. Nowcasting Indian GDP. *Oxford Bulletin of Economics and Statistics*, 80(2):259–282, 2018.

A. Caruso. Nowcasting with the help of foreign indicators: The case of mexico. *Economic Modelling*, 69:160–168, 2018. ISSN 0264-9993.

Crisil. New WPI, IIP series, Crisil first cut. *CRISIL Report*, May, 2017.

T. Dahlhaus, J. D. Guénette, and G Vasishtha. Nowcasting BRIC+M in real time. *International Journal of Forecasting*, 33(4):915–935, 2017.

C. De Mol, D. Giannone, and L. Reichlin. Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components? *Journal of Econometrics*, 146(2):318–328, 2008.

F. Dias, M. Pinheiro, and A. Rua. Forecasting portuguese gdp with factor models: Pre- and post-crisis evidence. *Economic Modelling*, 44:266–272, 2015. ISSN 0264-9993.

Francis X. Diebold. Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of diebold–mariano tests. *Journal of Business & Economic Statistics*, 33(1):1–1, 2015. doi: 10.1080/07350015.2014.983236.

Francis X Diebold and Roberto S Mariano. Com paring predictive accu racy. *Journal of Business and Economic Statistics*, 13(3):253–263, 1995.

T. Doan, R. Litterman, and C. Sims. Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews*, 3(1):1–100, 1984.

S. Ghosh and A. Ranjan. A machine learning approach to GDP nowcasting: An emerging market experience. *Bulletin of Monetary Economics and Banking*, 26:33–54, 2023.

D. Giannone, L. Reichlin, and L. Sala. Monetary policy in real time. *NBER Macroeconomics Annual*, 19:161–200, 2004.

D. Giannone, L. Reichlin, and D. Small. Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55(4):665–676, 2008. ISSN 0304-3932.

D. Giannone, S. M. Agrippino, and M. Modugno. Nowcasting china real gdp. *Technical report, Mimeo*, 2013.

T. Iyer and A. S. Gupta. Quarterly forecasting model for india's economic growth: Bayesian vector autoregression approach. *Asian Development Bank Economics Working Paper Series*, page 573, 2019a.

T. Iyer and A. S. Gupta. Nowcasting economic growth in India: The role of rainfall. *Asian Development Bank Economics Working Paper Series*, page 593, 2019b.

Malte Jahn. Artificial neural network regression models in a panel setting: Predicting economic growth. *Economic Modelling*, 91:148–154, 2020.

Y. Jiang, Y. Guo, and Y. Zhang. Forecasting china's gdp growth using dynamic factors and mixed-frequency data. *Economic Modelling*, 66:132–138, 2017. ISSN 0264-9993.

K. Kaustubh, S. S. Bhadury, and S. Ghosh. Reinvigorating GVA nowcasting in the post-pandemic period: A Case Study for India. *Bulletin of Monetary Economics and Banking*, 27:95–130, 2024.

P. Liu, T. Matheson, and R. Romeu. Real-time forecasts of economic activity for latin american economies. *Economic Modelling*, 29(4):1090–1098, 2012. ISSN 0264-9993.

R. S. Mariano and Y. Murasawa. A new coincident index of business cycles based on monthly and quarterly series. *Journal of applied Econometrics*, 18(4):427–443, 2003.

A. Richardson, v. F. M. Thomas, and T. Vehbi. Nowcasting gdp using machine-learning algorithms: A real-time assessment. *International Journal of Forecasting*, 37(2):941–948, 2021. ISSN 0169-2070.

M. Rusnák. Nowcasting czech gdp in real time. *Economic Modelling*, 54:26–39, 2016. ISSN 0264-9993.

T. J. Sargent and C. A. Sims. Business cycle modeling without pretending to have too much a priori economic theory. *New Methods in Business Cycle Research*, 1:145–168, 1977.

R. Sengupta. Some problems with india's gdp data and the way forward rajeswari sengupta. 2022. URL `https://api.semanticscholar.org/CorpusID:248914120`.

C. A. Sims. Macroeconomics and reality. *Econometrica: Journal of the Econometric Society*, 48(1):1–48, 1980.

J. H. Stock and M. W. Watson. New indexes of coincident and leading economic indicators. *NBER Macroeconomics Annual*, 4:351–394, 1989.

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

M. W. Watson. Monetary policy in real time: Comment. *NBER Macroeconomics Annual*, 19:216–221, 2004.

Q. Zhang, H. Ni, and H. Xu. Nowcasting chinese gdp in a data-rich environment: Lessons from machine learning algorithms. *Economic Modelling*, 122:106204, 2023. ISSN 0264-9993.

F. Özkan. Comparing the forecasting performance of neural network and purchasing power parity: The case of turkey. *Economic Modelling*, 31:752–758, 2013.

A Appendix

Table 9: Data Source and Availability

Series Name	Frequency	Source	Sample Start
Domestic Air Cargo	Monthly	Airports Authority of India	Dec-05
Auto Sales	Monthly	Society of Indian Automobiles Manufacturers	Dec-05
IIP Consumption	Monthly	Ministry of Statistics and Programme Implementation	Dec-05
Nong Import	Monthly	Ministry of Commerce and Industry	Dec-05
Rail Freight	Monthly	Indian Railways	Dec-05
Goods Export	Monthly	Ministry of Commerce and Industry	Dec-05
Foreign Tourists	Monthly	Ministry of Tourism	Dec-05
IIP Core	Monthly	Ministry of Statistics and Programme Implementation	Dec-05
Domestic Air Passengers	Monthly	Director General of Civil Aviation	Dec-05
IIP Cement	Monthly	Ministry of Statistics and Programme Implementation	Dec-05
IIP Electricity	Monthly	Ministry of Statistics and Programme Implementation	Dec-05
IIP Manufacturing	Monthly	Ministry of Statistics and Programme Implementation	Dec-05
Rail Passengers	Monthly	Indian Railways	Dec-05
PMI Manufacturing	Monthly	S&P Global	Jan-06

Continued on next page

Table 9: Data Source and Availability (Continued)

Series Name	Frequency	Source	Sample Start
PMI Services	Monthly	S&P Global	Jan-06
Crude Steel Consumption	Monthly	Joint Plant Committee	Dec-05
Petrol Consumption	Monthly	Petroleum Planning and Analysis Cell	Dec-05
Vehicle Registration	Monthly	Ministry of Roads and Highways	Jan-16
GST Revenue	Monthly	Goods and Services Tax Network	Oct-17
Sea Cargo	Monthly	Indian Ports Association	Apr-06
IIP Capital Goods	Monthly	Ministry of Statistics and Programme Implementation	Dec-05
Tractor Sales	Monthly	Tractor and Mechanisation Association	Jan-09
Electricity Demand	Monthly	Central Electricity Authority	Dec-05
Stringency Index	Monthly	University of Oxford	Dec-05
Ratio CPI to WPI	Monthly	Ministry of Statistics and Programme Implementation	Dec-05
Baltic Dry Index	Monthly	Baltic Exchange Information Services Limited	Dec-05
IMF Commodity Price Index	Monthly	International Monetary Fund	Dec-05
Naukri Jobspeak	Monthly	Naukri.com	Jul-08
CMIE Urban Unemployment	Monthly	Centre for Monitoring Indian Economy	Jan-16
New EPFO Subscribers	Monthly	Employees' Provident Fund Organisation	Sep-17
CMIE Rural Unemployment	Monthly	Centre for Monitoring Indian Economy	Jan-16

Continued on next page

Table 9: Data Source and Availability (Continued)

Series Name	Frequency	Source	Sample Start
E-Way Bills	Monthly	Goods and Services Tax Network	Apr-18
Hotel Occupancy Rate	Monthly	The Federation of Hotel & Restaurant Association of India	Apr-18
Average Revenue Per Available Room	Monthly	The Federation of Hotel & Restaurant Association of India	Apr-18
Gross Domestic Product	Quarterly	Ministry of Statistics and Programme Implementation	Dec-05
Gross Valus Added	Quarterly	Ministry of Statistics and Programme Implementation	Dec-05

Figure 6: Factors over Time (Transformed)

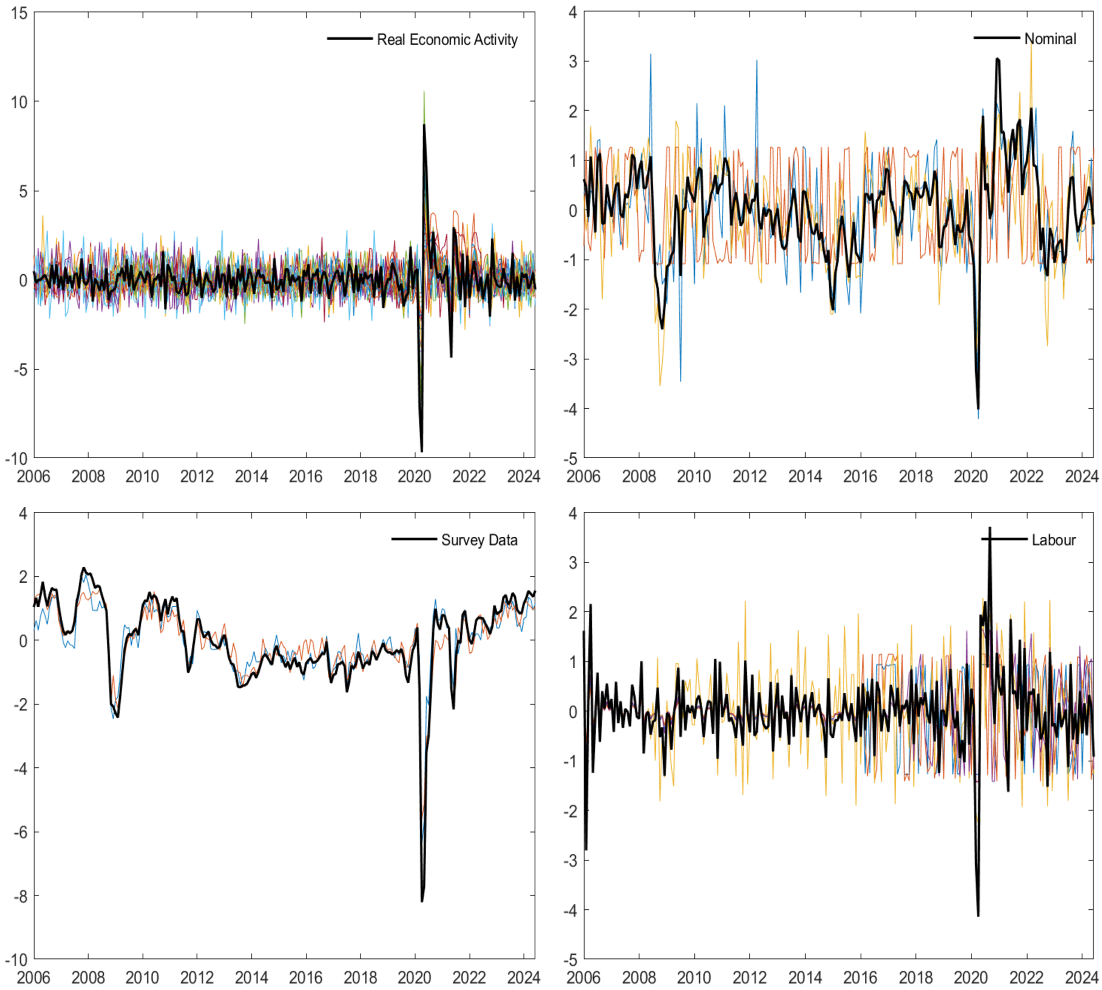


Figure 7: Comparison of movement of QoQ momentum of GVA and GDP with their projected QoQ momentum, based on factors (Transformed)