Digital Public Infrastructure and Bayesian Nowcasting of India's GDP: Why a Few Timely Indicators Suffice

Tandley Omprakash Sridevi*
Bharathidasan Institute of Management, Tiruchirappalli, India

Karan Singh Bagavathinathan
Independent Economist

This version: July 13, 2026

*Corresponding author. E-mail: sridevi.tandley@bim.edu

Abstract

A GDP nowcast is only as good as the small set of indicators it is built from, and in a rapidly digitising economy the craft of choosing them is being remade by newly available digital data. We update that craft for India, nowcasting GDP with the country's post-2017 digital public infrastructure, the Unified Payments Interface (UPI) and Goods and Services Tax (GST) records, alongside conventional official, administrative, and satellite series, within a Bayesian state-space ensemble. The digital layer merits inclusion for two reasons beyond its speed: it prints within days of the reference period, and it makes the large informal sector, invisible to traditional high-frequency indicators, partly observable. Guided by a bias–variance framework in which the accuracy-maximising indicator count rises with the sample length and with the number of orthogonal signals and falls with volatility, we find that out-of-sample accuracy is non-monotone: a compact set is optimal and richer specifications deteriorate. A five-indicator traditional workhorse (RMSE 0.74) and the preferred digital-augmented specification (0.72) each beat the Reserve Bank of India's Survey of Professional Forecasters (1.03) by a quarter to a third. We are candid that the digital layer's incremental gain is small and not yet significant ($p = 0.18$): at this early stage its variation is dominated by an adoption trend rather than the business cycle. The framework predicts that as adoption settles this trend recedes, the digital series' cyclical signal sharpens, and a few crafted digital indicators come to replace the traditional workhorse, delivering the same nowcast weeks earlier, a payoff that is scheduled rather than present and pre-registered for re-testing in 2028 and 2032.

Keywords: GDP nowcasting; digital public infrastructure; Bayesian state-space models; high-frequency data; India.

JEL classification: C53, C11, C32, E17, E37, O11, O47.

1 Introduction

How many indicators should a nowcast use? The dominant answer, formed on long samples of advanced economies, is effectively "as many as are available": with hundreds of series and decades of data, factor and shrinkage methods extract a common state and the marginal indicator rarely hurts (Stock and Watson, 2002; Bańbura and Modugno, 2014). We argue that this answer is not universal but conditional on an economy's informational maturity, and that in a young, data-scarce, rapidly digitising economy the accuracy-maximising indicator set is small and identifiable. We organise this claim with a bias–variance account in which the optimal indicator count k^* trades falling approximation bias against rising estimation variance and is increasing in the effective sample length and in the number of orthogonal signal dimensions and decreasing in macroeconomic volatility (Section 5); the account's general form and its cross-country test are developed in a companion study (Bagavathinathan and Sridevi, 2026), and we apply it where the tension is sharpest: India, whose national accounts combine a short reliable quarterly history with a large informal sector and a new, high-frequency digital data layer. The framework organises an otherwise puzzling empirical pattern (accuracy that improves and then sharply reverses as indicators are added), delivers a falsifiable prediction (parsimony should relax as the sample lengthens), and, as a by-product, produces a nowcast that outperforms the professional-forecaster consensus.

Macroeconomic policy in India rests on inferences about a latent state (output, inflation pressure, capacity use) that is never directly observed and is recovered by combining noisy, lagged indicators within a state-space framework (Harvey, 1989; Giannone et al., 2008). The exercise faces two challenges that compound in emerging economies. First, the informal sector (over four-fifths of employment in India) is tracked only through periodic surveys that themselves arrive with release lags. Second, the high-frequency indicators that do exist, principally the Index of Industrial Production, cover predominantly formal manufacturing, leaving services and small-trade activity invisible at high frequency.

A sequence of reforms over the past decade has begun to close both gaps. The Goods and Services Tax (2017) brought small enterprises into formal tax registration and now yields monthly

returns. The bulk of nominally informal firms remain outside the GST system (the 2023-24 Annual Survey of Unincorporated Sector Enterprises (ASUSE) records GST registration at 2.1 percent of the 73.4 million surveyed establishments), but participation has been rising gradually as compulsory-registration thresholds capture a growing tail of small firms. The Unified Payments Interface (2016) processes a large and growing volume of small-value transactions; Crouzet et al. (2023) establish empirically that demonetisation accelerated UPI adoption among small Indian merchants who previously transacted in cash. Information-and-communications technology penetration among informal enterprises is also expanding rapidly: ASUSE finds that the share of unincorporated firms using the internet for entrepreneurial purposes nearly doubled from 13.9 percent in 2021-22 to 26.7 percent in 2023-24, and the 2026 ASUSE round currently in the field will for the first time measure direct use of UPI, online banking, point-of-sale devices, and payment gateways among informal enterprises (Ministry of Statistics and Programme Implementation, 2025). The Jan Dhan financial-inclusion programme and the digitisation of utility billing have produced near-continuous administrative records of credit, savings, and energy use. Collectively this digital public infrastructure (DPI) generates high-frequency records of activity that previously appeared in official statistics with substantial lag, in aggregated form, or not at all; coverage will deepen as digital adoption continues to spread, and the framework's effective informal-sector visibility improves with each successive survey round.

A second segment sits outside ASUSE's establishment frame by construction: platform-mediated gig work: ride-hailing, food delivery, and personal services. Under Section 9(5) of the GST Act the e-commerce operator bears GST liability on these notified services regardless of the worker's registration status, and the worker-side compensation flows through UPI, so the same digital rails are the only administrative channels through which the segment is currently observable. NITI Aayog (2022) estimate 7.7 million such workers in 2020–21, projected to reach 23.5 million by 2029–30, a fast-growing informal slice that no establishment survey can reach, but that GST collections and UPI transaction volume can.

Whether this new layer of data carries signal beyond what official statistics already capture is

not obvious. GST collections co-move with manufacturing output; UPI volumes co-move with retail consumption. The question is therefore empirical: do high-frequency administrative and transactional indicators reduce nowcast error when added to a baseline that uses only official statistics, and if so, by how much given that the new data are noisy and short? We answer this question for India and discuss what the answer implies for the framework's evolving statistical power as the post-2017 sample lengthens.

We construct a quarterly panel of thirteen indicators covering 2019Q3 through 2026Q1 and fit seven progressive specifications M_1 through M_7 that incrementally add data-source groups (official, administrative, satellite, transactional, textual, and global) combining Ridge regression, gradient boosting, and, in the final two specifications, a Bayesian linear-Gaussian state-space layer. Out-of-sample evaluation uses an expanding window over 2023Q2 to 2026Q1, with the Clark–West (2007) test as the headline nested-model significance metric.

The specification that adds UPI transaction volume to the official, administrative, and satellite indicators (M_4) attains the lowest out-of-sample RMSE of 0.72 percentage points, improving on the official-only baseline (M_1, RMSE 0.77) by 7.5 percent. The Clark–West statistic is in the predicted direction (+0.97) but does not reach conventional significance at the twelve-quarter sample length (one-sided $p = 0.18$). We interpret this as a power-limited result rather than evidence against the framework: the comparison would require either a larger improvement or a longer sample to discriminate at conventional levels. Consistent with this reading, a variance decomposition attributes 73 percent of forecast uncertainty to directly reducible components (epistemic plus specification) and a further 24 percent to partially reducible policy uncertainty, indicating that the framework's discriminating power is limited primarily by sample length rather than by intrinsic noise of the indicators.

This paper sits at the intersection of three literatures (state-space nowcasting, alternative-data macro measurement, and informal-sector measurement) that have not previously met. We draw on each: methodology from the state-space tradition, an indicator philosophy from the alternative-data literature extended to India's post-2017 digital public infrastructure, and motivation from the

informal-sector measurement literature. Section 2 develops the contribution in detail. Section 3 presents the data, Section 4 the model, Section 6 the empirical results, and Section 7 discusses the policy implications.

2 Literature Review

2.1 Nowcasting and state-space methods

The state-space representation treats observed indicators as noisy measurements of an unobserved latent state evolving under its own dynamics (Harvey, 1989; Durbin and Koopman, 2012); the Kalman filter estimates that state recursively from any observed history, making the framework natural when data arrive at irregular frequencies, and it has a track record in Indian macroeconomic measurement (Singh et al., 2011). It has since been extended to mixed-frequency GDP nowcasting, news-decomposition of release-by-release revisions, and cross-sectionally dependent panels (Giannone et al., 2008; Bańbura and Modugno, 2014; Fosten and Nandi, 2023), and now anchors the nowcasts of the Federal Reserve Bank of New York and the European Central Bank. Bayesian implementations add two advantages for policy use: parameter uncertainty is propagated into the forecast density rather than hidden behind point estimates (Geweke and Amisano, 2010), and priors discipline latent dynamics when the sample is short relative to the parameter count, a recurring constraint in Indian quarterly data, where releases begin in 1996 and breaks at demonetisation, GST, and COVID further reduce effective sample size.

For India, dynamic-factor and coincident-indicator nowcasts, Bayesian VARs, and RBI's Economic Activity Index all draw on conventional monthly indicators and document the large revisions between first and final GDP releases (Bragoli and Fosten, 2018; Bhadury et al., 2019; Kumar, 2020; Bhattacharya et al., 2023; Iyer and Gupta, 2019; Qureshi et al., 2026). The most recent, Kaustubh and Ranjan (2025), consolidates high-frequency series into a mixed-frequency VAR but uses reform-era data such as GST revenue and e-way bills only at aggregate level, excluding the transactional channels we exploit.

2.2 High-frequency and alternative data

A growing literature shows that alternative data (from globally consistent satellite night-light intensity to transactional series such as credit-card, payroll, and small-business records embedded in mixed-frequency models) can track economic activity in near real time, with predictive power strongest where official statistics are weak (Henderson et al., 2012; Chetty et al., 2020; Eckert et al., 2025). Such indicators are powerful but agnostic about sectoral composition or transaction channels, and are largely private-sector by-products released at their owners' discretion.

India's reforms have produced a distinct fourth class, administrative, transaction-level, and policy-mandated data we term digital public infrastructure (DPI): the GST (2017) and UPI (2016), with Aadhaar-linked bank accounts under Jan Dhan, utility-billing digitisation, and the e-Way bill system. D'Silva et al. (2019) characterise this "India Stack" as deliberate institutional design rather than spontaneous innovation; Carrière-Swallow et al. (2021) document UPI's implications for financial inclusion; and Crouzet et al. (2023) show demonetisation accelerated digital-payment adoption among small merchants previously invisible to formal data. DPI is thus administratively bounded to one jurisdiction but comprehensive within it, and carries a paper trail connecting to formal-economy registries. Beyond payments, this infrastructure leaves a measurable economic footprint at the micro level: digital financial inclusion smooths households' transitory consumption in India (Kumar et al., 2026), and ICT adoption raises enterprise productivity (Garg et al., 2026), channels through which DPI-linked activity feeds the very macroeconomy it helps to measure.

2.3 Informal-sector measurement

A separate literature measures informal activity, which standard national-accounts apparatus systematically under-records. La Porta and Shleifer (2014) document that informality often exceeds 40 percent of measured GDP in low- and middle-income economies, and Schneider et al. (2010) infer shadow-economy size indirectly through currency-demand and MIMIC methods; by definition, informal activity leaves no firm-survey or tax-return footprint at the time of transaction. Indian modelling has long worked around this constraint, relying on partial proxies (the Index of Industrial

Production for output and the CPI for Industrial Workers for inflation) whose mismeasurement complicates empirical macro work such as Phillips-curve estimation (Singh et al., 2011); the exclusion of informality was practical, not principled.

Recent improvements (the composite CPI (2014), the quarterly Periodic Labour Force Survey, the annual Survey of Unincorporated Sector Enterprises from 2021–22, and RBI's 27-indicator Economic Activity Index (Kumar, 2020)) still deliver no real-time reading on informality: the surveys remain quarterly or annual, and the EAI's indicator set predates the post-2017 DPI buildout. A natural objection is that formal-sector indicators capture informal activity through supply-chain linkages, but this holds only as a long-run statement and has not been tested at nowcasting frequency; reform-era integration of small firms (through GST registration thresholds and UPI adoption (Crouzet et al., 2023)) is itself shifting those linkages, so the channel that supposedly carries informal information through formal proxies is neither stable nor independently estimated.

2.4 Research gap and contribution

Three observations follow, consistent with the systematic survey of Stundziene et al. (2024) across 193 nowcasting studies. First, the nowcasting literature applies the state-space framework primarily to formal-sector indicators; even recent Indian work (Kaustubh and Ranjan, 2025) uses reform-era series only at aggregate level. Second, the alternative-data literature has broadened the indicator set but remains centred on advanced economies where informality is small. Third, the informal-sector literature offers indirect estimates and surveys but no real-time, transaction-level view. The three strands have not met.

We draw on each: methodology from the state-space tradition, an indicator philosophy from the alternative-data literature extended to India's post-2017 DPI, and motivation from informal-sector measurement. Three features distinguish our contribution. First, the indicator set incorporates the post-2017 transactional layer (UPI volume, GST collections) alongside conventional administrative and satellite series, giving a high-frequency window onto partially informal activity, though UPI volume proxies digital-payment penetration rather than directly measuring the informal sec-

tor. Second, the model is genuinely Bayesian: parameter uncertainty is propagated through the Kalman-filter posterior of a linear-Gaussian state-space model estimated by EM (Shumway and Stoffer, 1982). Third, we adopt an honest small-sample evaluation: nested-model comparison uses the Clark–West (Clark and West, 2007) test, with the statistic and one-sided p-value reported transparently rather than claimed as significant at twelve out-of-sample observations. Section 6 reports that the specification adding UPI volume attains the lowest out-of-sample RMSE, with the gain in the predicted direction but not significant at the present sample length.

3 Data

This section describes the data used in the empirical analysis, including sources, construction of variables, and key features of the sample.

3.1 Data sources

The dataset is a quarterly panel covering 2019Q3–2026Q1 (27 quarters), bounded below by the post-2017 emergence of GST (July 2017) and UPI (meaningful volume from 2017Q3); 2019Q3 is the first quarter for which all series have a complete year-on-year back-history for transformation.

The dependent variable is real GDP growth from MoSPI's new series on a 2022–23 base. Domestic high-frequency indicators comprise three groups. *Official indicators*, IIP and headline consumer-price inflation (MoSPI), are the conventional inputs to India's nowcasting literature. *Administrative indicators* (GST collections (GSTN, Ministry of Finance), non-food bank credit (RBI), and electricity supplied (Central Electricity Authority)) capture activity flows produced as a by-product of policy-mandated reporting and physical-flow measurement. *Transactional indicators*, UPI transaction volume (NPCI), capture the rapidly digitising informal economy; the Indian payments system now routes a substantial share of small-value transactions previously invisible to official statistics. Two further sources enter the model: satellite night-time lights (NASA/NOAA Black Marble VIIRS) and the Economic Policy Uncertainty index for India (Baker et al., 2016).

The model also includes four global controls (Brent crude (FRED/EIA), the Geopolitical Risk index (Caldara and Iacoviello, 2022), the INR/USD reference rate (RBI), and the effective federal funds rate (Federal Reserve)) entered as stationary transforms of their level series.

Four measurement notes apply to recent observations. First, non-food bank credit replaces total bank credit (food credit is below 0.5 percent of total and the non-food measure is the standard indicator for non-agricultural activity). Second, GST collections are measured as gross collections excluding the compensation cess (the sum of CGST, SGST, and IGST components): the GSTN's new-format statistics from FY 2025–26 onward no longer report the cess (roughly 7 percent of collections), so the ex-cess definition is the only basis consistent across the sample; it is also the basis on which the official monthly press releases compute growth rates. March 2026 collections are taken from the official monthly press release, as the statewise workbook extends only through February 2026. Third, CPI for January–March 2026 uses MoSPI's new 2024=100 base (the 2012=100 series ends December 2025); the quarterly year-on-year rate is computed entirely within the new-base series, so no linking factor is required. Fourth, the INR/USD reference rate uses RBI publication through July 2018, the FBIL reference rate from April 2022, and FBIL bridging values for the August 2018–March 2022 RBI publication gap.

3.2 Data sources and publication lag

Table 1 summarises each series by source group, native frequency, and publication lag. Official statistics carry the longest release lags (GDP at 45–60 days, IIP at 30–45 days), administrative records are available within two to five weeks of the reference period, and transactional and textual series within days.

3.3 Variable construction

Series subject to within-year seasonality enter the model as year-on-year growth rates (GDP, GST collections, UPI volume, IIP, non-food bank credit, electricity supplied, CPI inflation), which removes seasonal variation by construction; we do not apply a separate seasonal-adjustment pro-

Table 1: Data sources, source group, native frequency, and publication lag

Variable	Source Group	Source	Frequency	Lag (days)
GDP	Official	MoSPI (Base 2022-23)	Quarterly	45–60
IIP	Official	MoSPI	Monthly	30–45
CPI Inflation	Official	MoSPI	Monthly	10–15
GST Collections	Administrative	GSTN / MoF	Monthly	15–20
Non-Food Bank Credit	Administrative	RBI	Fortnightly	30–35
Power Consumption	Administrative	CEA	Monthly	20–25
UPI Transactions	Transaction	NPCI	Daily	1–2
Night Lights (VIIRS)	Satellite	NASA/NOAA Black Marble	Daily	2–5
Policy Uncertainty (EPU)	Textual	Baker et al. (2016)	Monthly	0–1
Brent Crude	Global	FRED / EIA	Daily	0–1
GPR Index	Global	Caldara and Iacoviello (2022)	Daily	0–5
INR/USD Rate	Global	RBI / FBIL	Daily	0–1
Fed Funds Rate	Global	Federal Reserve	Monthly	0–1

Notes: All variables are aggregated to calendar quarterly frequency. Estimation sample 2019Q3–2026Q1 (N=27). Quarter labels are calendar quarters (e.g. 2020Q2 = April–June 2020); India's official GDP series is published on a fiscal-year basis (April–March). GST collections are gross excluding compensation cess (see measurement notes).

cedure. Night-light intensity enters as log(1 + radiance). Global controls enter as stationary transforms of their level series: Brent crude, the INR/USD rate, and the GPR index as year-on-year percentage changes; the federal funds rate as quarter-on-quarter first-difference. Two lags of GDP growth are included as features to absorb residual autocorrelation in the target.

3.4 Descriptive statistics

Table 2 reports descriptive statistics for the 13 indicators, each in the form used by the model; all are fully observed ($N = 27$). GDP growth is strongly left-skewed (skewness -1.90, excess kurtosis 7.67), driven by the COVID contraction (2020Q2, -23.1%) and rebound (2021Q2, $+22.6\%$), both contained within the training window.

3.5 Stationarity

Stationarity is assessed using both the augmented Dickey–Fuller (ADF) test, with null hypothesis of a unit root, and the Kwiatkowski–Phillips–Schmidt–Shin (Kwiatkowski et al., 1992) test, with

Table 2: Descriptive statistics, 2019Q3–2026Q1 ($N = 27$)

Variable	Mean	SD	Skew	Kurt	Min	Max
GDP Growth (%)	5.50	7.35	−1.90	7.67	−23.13	22.63
GST Collection (YoY %)	12.59	19.01	0.87	5.70	−40.96	79.03
UPI Volume (YoY %)	73.37	35.76	0.80	0.28	25.94	171.84
IIP Growth (YoY %)	3.56	11.71	0.19	7.92	−35.56	44.39
CPI Inflation (YoY %)	5.10	1.64	−0.96	0.35	0.76	7.28
Non-Food Credit Growth (YoY %)	11.74	5.03	0.30	−1.24	5.11	20.17
Power Consumption (YoY %)	4.43	6.80	−0.64	1.79	−16.20	17.65
Policy Uncertainty Index	103.75	44.70	1.41	2.74	45.21	254.35
Night Light (log)	0.29	0.24	−0.48	−0.89	−0.20	0.60
Brent Crude (YoY %)	7.09	42.36	1.28	1.21	−56.98	132.27
INR/USD Depreciation (YoY %)	3.43	3.22	0.31	−0.52	−2.80	9.67
GPR Index (YoY change)	15.40	46.66	1.93	5.21	−50.80	189.02
Fed Funds Change (QoQ pp)	0.05	0.56	0.80	1.42	−1.20	1.46

Notes: Sample 2019Q3–2026Q1 ($N = 27$); all variables fully observed. Each row shows the indicator in the form used by the model: year-on-year growth for series with seasonal structure (GDP, GST, UPI, IIP, CPI, non-food credit, power, Brent, INR/USD, GPR), level for the EPU index, log-radiance for night lights, and first-difference for the federal funds rate. Skew = sample skewness; Kurt = excess kurtosis. The COVID quarters 2020Q2–2021Q1 fall within the estimation sample.

null hypothesis of stationarity, reported in Table 3. Classification of integration order in the *Order* column uses the ADF result; KPSS statistics are reported as a robustness check.

Of the 13 series tested in transformed form, 4 are classified $I(0)$, 6 as $I(1)$, and 3 as $I(2)$? (non-food credit growth, the EPU index, and log night lights). ADF and KPSS conflict for 6 of the 13 series; we read this as the limited power of unit-root tests at $N = 27$ rather than non-stationarity, since the year-on-year and first-difference transformations used by the model are stationary by construction.

4 Empirical Framework and Methodology

This section sets out the model specifications, the component learners and their ensemble combination, and the out-of-sample evaluation procedure. We build the forecasting framework as a sequence of nested specifications $M_1, \ldots, M_7$, each adding one data source group, so that incremental gains

Table 3: Unit root tests (ADF with KPSS as robustness)

Variable	ADF	ADF p	KPSS	KPSS p	ΔADF	ΔADF p	Order
GDP Growth (%)	−5.607	0.000	0.276	0.100	−3.392	0.011	$I(0)$
GST Collection (YoY %)	−5.037	0.000	0.124	0.100	−3.607	0.006	$I(0)$
UPI Volume (YoY %)	−0.992	0.756	0.689	0.015	−2.901	0.045	$I(1)$
IIP Growth (YoY %)	−1.465	0.551	0.139	0.100	−3.098	0.027	$I(1)$
CPI Inflation (YoY %)	−1.818	0.371	0.453	0.054	−5.600	0.000	$I(1)$
Non-Food Credit Growth (YoY %)	−0.995	0.755	0.416	0.070	−2.738	0.068	$I(2)$?
Power Consumption (YoY %)	−3.446	0.009	0.185	0.100	−3.885	0.002	$I(0)$
Policy Uncertainty Index	−0.240	0.934	0.548	0.031	−0.078	0.952	$I(2)$?
Night Light (log)	−2.373	0.149	0.746	0.010	−1.746	0.407	$I(2)$?
Brent Crude (YoY %)	−2.236	0.193	0.131	0.100	−4.859	0.000	$I(1)$
INR/USD Depreciation (YoY %)	−2.011	0.282	0.091	0.100	−3.702	0.004	$I(1)$
GPR Index (YoY change)	−3.491	0.008	0.092	0.100	−5.913	0.000	$I(0)$
Fed Funds Change (QoQ pp)	−2.198	0.207	0.160	0.100	−3.436	0.010	$I(1)$

Notes: Augmented Dickey–Fuller (ADF) test with null of a unit root (5% critical value ≈ -3.0); Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test with null of stationarity. Maximum lag set by the Schwert (1989) rule. *Order* uses ADF only: $I(0)$ if ADF rejects at level; $I(1)$ if it rejects at first difference but not at level; $I(2)$? if it fails at both. ADF and KPSS conflict for 6 of 13 series; given $N = 27$ both tests have limited power, and we read conflict as low power rather than non-stationarity. All year-on-year transformations are stationary by construction.

can be attributed to specific indicator types. Models M_1 through M_5 combine penalised linear regression with gradient boosting; the final two specifications add a Bayesian linear-Gaussian state-space model. We deliberately favour interpretable, well-identified components over deep non-linear architectures: with 15 training quarters and 12 out-of-sample quarters, complex models have insufficient data to discipline parameter estimates, and small-sample power is the binding constraint.

4.1 Progressive model specifications

Let y_t denote real GDP growth in quarter t, and let $\mathbf{x}_t$ collect the contemporaneous indicators described in Section 3. Each specification predicts y_t from $(\mathbf{x}_t, y_{t-1}, y_{t-2})$ using an incrementally larger feature set. The seven specifications are:

- M_1 *(Official Only)*: GDP lags, IIP, CPI inflation.
- M_2 *(+ Administrative)*: M_1 + GST, non-food bank credit, electricity supplied.

- M_3 *(+ Satellite)*: M_2 + night-light radiance.
- M_4 *(+ Transaction)*: M_3 + UPI volume.
- M_5 *(+ Textual)*: M_4 + Economic Policy Uncertainty.
- M_6 *(+ Global)*: M_5 + Brent crude, INR/USD, GPR, federal funds.
- M_7 *(Full BSSM)*: same feature set as M_6 with a Bayesian state-space component added to the ensemble.

The nesting structure permits the use of the Clark–West (Clark and West, 2007) test for nested-model forecast comparison: $M_2, \ldots, M_7$ each strictly contain M_1, so under the null that the larger model adds no signal, the larger model's MSPE is mechanically inflated by parameter-estimation noise; the CW correction adjusts for this bias.

All indicators enter the model at quarterly frequency: native-daily series (UPI volume, night-light radiance, Brent, INR/USD, GPR) are aggregated to quarterly sums or end-of-quarter snapshots; native-monthly series (IIP, CPI, GST, electricity, EPU, federal funds) are quarterly-averaged; native-fortnightly bank credit is taken at end-of-quarter. Exploiting within-quarter timing through a mixed-frequency Kalman filter (Giannone et al., 2008) is a natural extension we leave for future work given the short post-2017 sample.

4.2 Component learners: Ridge and Gradient Boosting

Each specification produces a forecast as an inverse-variance weighted average of two base learners, penalised linear regression (Ridge) and gradient boosting (GBR), fitted on the same feature set. The two learners capture complementary structure: Ridge (Hoerl and Kennard, 1970) enforces a linear, shrinkage-regularised mapping that is robust at small N; GBR (Friedman, 2001) captures any non-linear interactions present in the small training window without imposing a parametric form.

Let $\mathbf{z}_t$ denote the standardised feature vector at quarter t. The Ridge component solves

$$\hat{\boldsymbol{\beta}}_R = \arg\min_{\boldsymbol{\beta}} \sum_{s=1}^{t-1} \left(y_s - \mathbf{z}_s'\boldsymbol{\beta}\right)^2 + \alpha\|\boldsymbol{\beta}\|_2^2, \tag{1}$$

with regularisation parameter $\alpha = 5$ held fixed across all specifications. The GBR component fits 80 regression trees of maximum depth 3 with learning rate 0.05; hyperparameters are held fixed across specifications so that differences in performance reflect data content rather than tuning. Both learners produce a point forecast and an in-sample residual standard deviation; the latter is inflated by $\sqrt{1 + 1/t}$ to obtain an out-of-sample-consistent prediction standard error.

4.3 Bayesian state-space component

Specifications M_6 and M_7 add a Bayesian linear-Gaussian state-space model (BSSM) as a third ensemble component. The BSSM captures the contemporaneous co-movement of all indicators with GDP growth through a low-dimensional latent factor, providing structural advantage over the autoregressive base learners for current-quarter nowcasting.

The state-space representation is

$$\mathbf{z}_t = F\mathbf{z}_{t-1} + G\mathbf{u}_t + \mathbf{w}_t, \qquad \mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, Q), \tag{2}$$

$$\mathbf{y}_t = H\mathbf{z}_t + \mathbf{v}_t, \qquad \mathbf{v}_t \sim \mathcal{N}(\mathbf{0}, R), \tag{3}$$

where $\mathbf{z}_t \in \mathbb{R}^n$ is a latent economic-activity state ($n = 2$ for M_6, $n = 3$ for M_7), $\mathbf{u}_t = (y_{t-1}, y_{t-2})'$ is an exogenous input vector of lagged GDP, and $\mathbf{y}_t$ is the observation vector of contemporaneous indicators (including GDP, which is masked at test time). The transition matrix F has its spectral radius clipped to 0.97 to ensure stationarity. Parameters $\{F, G, H, Q, R\}$ are estimated by EM (Shumway and Stoffer, 1982), with the E-step implemented as a Kalman filter followed by an RTS smoother. The observation loading H is warm-started by correlation with GDP to prevent the degenerate fixed point in which the GDP loading collapses to zero.

Two implementation details merit note. First, the Kalman filter handles missing values exactly: in the update step at each quarter, only rows of $\mathbf{y}_t$ that are observed contribute to the innovation. No imputation is required. Second, the BSSM activates in M_6, M_7 only when the training window has reached MIN TRAINING + 5 = 20 quarters; below this threshold the EM is insufficiently disciplined and the ensemble falls back to Ridge+GBR.

4.4 Ensemble combination

Within each specification, the component forecasts are combined by inverse-variance weighting (IVW):

$$w_i = \frac{1/\sigma_i^2}{\sum_j 1/\sigma_j^2}, \qquad \hat{y}_t = \sum_i w_i \, \hat{y}_{i,t}, \tag{4}$$

with a per-component floor of $w_i \geq 0.15$ to prevent any single learner from dominating when its training residual variance is fortuitously near zero. The combined predictive variance follows the additive form $\sigma_t^2 = \sum_i w_i^2 \sigma_{i,t}^2$, conservatively assuming uncorrelated component errors.

4.5 Out-of-sample evaluation

Forecast accuracy is evaluated using an expanding-window pseudo-out-of-sample procedure. At each quarter $t \in \{2023Q2, \ldots, 2026Q1\}$, the model is re-estimated on data through $t-1$ and used to predict y_t; the initial training window covers 15 of 27 quarters (2019Q3–2023Q1, including all four COVID quarters), leaving 12 quarters for out-of-sample evaluation. Hyperparameters are held fixed across specifications and across expanding-window iterations.

Three significance tests are reported. The Diebold–Mariano (Diebold and Mariano, 1995) test with the Harvey–Leybourne–Newbold (Harvey et al., 1997) small-sample correction tests whether the loss differential between two models has zero mean; we use the squared-error loss and a two-sided alternative. The Clark–West (Clark and West, 2007) test is the appropriate test for nested-model forecast comparison and is reported as the headline significance metric; the CW statistic is constructed under a one-sided alternative that the larger model dominates the baseline.

A CRPS-based variant of the DM test is reported for density forecast quality.

With $N = 12$ out-of-sample observations the tests have limited statistical power. We interpret the empirical results accordingly: the direction and magnitude of forecast-accuracy differences are informative, but the strict significance levels are conservative bounds at small N.

4.6 Forecast uncertainty decomposition

We decompose the predictive variance into four components to identify the dominant sources of forecast uncertainty:

$$\mathrm{Var}(y_t) = \sigma^2_{\text{alea}} + \sigma^2_{\text{epi}} + \sigma^2_{\text{spec}} + \sigma^2_{\text{policy}}. \tag{5}$$

The *aleatoric* component is the irreducible noise of the data-generating process, estimated from the BSSM observation-noise covariance R. The *epistemic* component captures parameter-estimation uncertainty and shrinks as the training sample grows; it is computed from the Kalman filter's posterior state covariance at the prediction quarter. The *model specification* component is the variance across the seven ensemble predictions, capturing structural uncertainty across feature configurations. The *policy uncertainty* component is the variance of the prediction across configurations that differ in EPU loading, isolating the policy-uncertainty channel. The decomposition makes precise which components shrink with additional data (epistemic, partially policy) and which do not (aleatoric).

5 A theory of indicator parsimony and economic maturity

The progressive specifications of Section 4 embody a hypothesis that we now make explicit: the out-of-sample accuracy of a nowcasting model is not monotone in the number of indicators, and the location of its optimum is governed by the economy's informational maturity. We set out the mechanism as the framework this paper applies to India, and extract the within-country prediction that becomes testable as the sample lengthens.

For a nowcast using k indicators, the companion study (Bagavathinathan and Sridevi, 2026)

writes the expected out-of-sample mean-squared forecast error as the sum of a squared-bias and an estimation-variance term,

$$\mathrm{MSFE}(k) \;\approx\; B^2(k) \;+\; \kappa^2 \,\frac{k}{N}, \tag{6}$$

where N is the effective training length and κ^2 scales the per-parameter estimation noise. The squared bias $B^2(k)$ falls as indicators are added and flattens once the r dominant common factors are spanned, since further indicators are largely collinear with those already in the model; the variance term rises linearly in k, as each retained indicator must be paid for with an estimated coefficient. Equation (6) is therefore U-shaped, with an interior optimum k^* where the marginal bias reduction equals the marginal estimation cost κ^2/N. We carry only as much of this framework as the India application requires.

The framework yields three comparative statics, established formally in the companion study: the optimal indicator count k^* is non-decreasing in the training length N, non-decreasing in the number of orthogonal signal dimensions r, and non-increasing in the per-parameter noise scale κ^2.

The three comparative statics map directly onto the stage of statistical and economic development. An emerging economy is characterised by a short reliable data history (small N, in India's case the transactional layer used here begins only in 2017), by a narrow and partially measured formal sector that yields few orthogonal signal dimensions (small r, as administrative and transaction series co-move strongly with one another and with output), and by elevated macroeconomic volatility from structural change and incomplete stabilisation (large κ^2). All three forces push k^* downward. A mature economy occupies the opposite corner (long samples, many well-measured and weakly correlated sectors, and a stable data-generating process) where k^* is large and the marginal indicator continues to help. The large-dataset nowcasting tradition (Stock and Watson, 2002; Bańbura and Modugno, 2014), developed almost entirely on long-sample advanced economies, is on this reading the high-N, high-r limit of a more general relationship rather than a universal prescription. The recurrent finding that simple, parsimonious combinations outperform richly parameterised "optimal" ones (the forecast-combination puzzle (Stock and Watson, 2004; Smith and Wallis, 2009), the "less-is-more" effect (Gigerenzer and Brighton, 2009), and the illusion

of sparsity in macroeconomic prediction (Giannone et al., 2021)) is the same U-shape viewed from the small-N side.

Two empirical implications follow. First, within a single economy k^* should rise as the sample lengthens: the parsimony that is optimal today should relax as N grows and the variance term in (6) shrinks. This yields a falsifiable, near pre-registered prediction: the re-estimation checkpoints adopted in Section 7 ($N = 22$, $\approx$2028; $N = 35$, $\approx$2032) should show the optimum migrating from its present location toward richer specifications, and a failure of k^* to rise would count as evidence against the mechanism. Second, the accuracy cost of over-inclusion should be larger in higher-volatility episodes, where κ^2 amplifies the variance term, a pattern consistent with the sub-period analysis reported in the Supplementary Material.

We are explicit about what these data can and cannot establish. With a single country we can state the mechanism, locate India in its small-N corner, and test the within-country prediction over time; we cannot test the cross-country comparative statics directly, which would require a panel of economies at different stages of informational maturity. We therefore offer Equation (6) and its proposition as a framework that organises the non-monotone accuracy pattern documented below (the optimum at M_4 and the degradation of M_5–M_7) rather than as a cross-country law; the panel test is the natural next step, which a companion cross-country study takes up directly, finding that the optimal indicator count tracks economic complexity across economies and places India at the low-complexity, few-indicator end (Bagavathinathan and Sridevi, 2026). That a compact official-and-administrative information set suffices to nowcast Indian GDP (with accuracy that ceases to improve, and eventually deteriorates, as further indicators are added) is precisely the behaviour Equation (6) predicts at the small-N, small-r pole. The present single-country evidence thus corroborates the parsimony mechanism from within, and locates India as one calibrated point on the informational-maturity gradient that the companion panel traces across economies.

6 Empirical Results

This section presents the empirical results in two parts. We begin with overall forecast accuracy across the seven specifications (Section 6.1), evaluated on the post-COVID out-of-sample window 2023Q2–2026Q1 ($N = 12$). We then decompose the predictive uncertainty into reducible and irreducible components (Section 6.2), which informs the discussion of how the framework's discriminating power evolves with sample length. Robustness exercises (leave-one-out source dropping, feature-count progression, and a sub-period disaggregation of the OOS window) are reported in the online Supplementary Material.

6.1 Forecast accuracy

Table 4 reports out-of-sample forecast accuracy in two panels. Panel A gives the seven crafted nested specifications, with Diebold–Mariano and Clark–West statistics relative to the official-only baseline M_1; Panel B reports a compact no-digital workhorse, discussed in the parsimony comparison below.

Three findings emerge from Table 4. *First, M_4 (the specification that adds UPI transaction volume to the official-and-administrative-and-satellite information set) attains the lowest out-of-sample RMSE of 0.72 percentage points*, improving on M_1's 0.77 by 7.5 percent. *Second, the gain is in the direction predicted but is not conventionally significant*: the Clark–West statistic of +0.97 corresponds to a one-sided p-value of 0.18. We interpret this as a power-limited result consistent with the 12-quarter evaluation window; the implications for the appropriate framing of the paper's contribution are discussed in Section 7. *Third, specifications M_5 through M_7 (which add textual, global, and state-space components) sharply underperform the baseline*, with RMSEs in the 2.1 to 2.8 range. With 15 training quarters and 10–15 features, M_5 through M_7 are over-parameterised, and the IVW ensemble does not adequately discipline parameter estimates at the corresponding feature dimensions. The non-monotone profile (accuracy improving to M_4 and then deteriorating) is the U-shape predicted by Equation (6): with $N = 15$ the estimation-variance term dominates beyond $k^* \approx 9$, placing India squarely in the small-N corner of Section 5 where parsimony is

Table 4: Less is more: forecast accuracy of crafted specifications versus a compact workhorse (out-of-sample 2023Q2–2026Q1, $N = 12$)

Model (k)	RMSE	MAE	MAPE (%)	CRPS	Cov 95%	DM	p(DM)	CW	p(CW)
Panel A. Crafted nested specifications									
M_1: Official only (4)	0.77	0.65	8.8	0.47	83.3	—	—	—	—
M_2: + Administrative (7)	0.86	0.71	9.6	0.53	66.7	−0.55	0.60	+0.13	0.45
M_3: + Satellite (8)	0.80	0.63	8.5	0.48	75.0	−0.14	0.89	+0.55	0.30
M_4: + Transaction (9)	**0.72**	**0.56**	**7.5**	**0.41**	75.0	+0.40	0.70	+0.97	**0.18**
M_5: + Textual (10)	2.07	1.18	16.3	1.00	58.3	−0.98	0.35	−0.69	0.75
M_6: + Global (15)	2.79	1.71	24.3	1.49	50.0	−1.28	0.23	−1.15	0.86
M_7: Full BSSM (15)	2.78	1.68	24.0	1.47	50.0	−1.27	0.23	−1.08	0.85
Panel B. Compact workhorse (traditional, no digital)									
Workhorse: traditional (5)	0.74	0.54	7.3	0.41	—	—	—	—	—
+ credit, power (7)	0.90	0.71	9.6	0.56	—	—	—	—	—
+ UPI, GST (= M_4; 9)	0.72	0.56	7.5	0.41	—	—	—	—	—
Memo: RBI SPF consensus	1.03	0.91	—	—	—	—	—	—	—

Notes: Out-of-sample window 2023Q2–2026Q1 ($N = 12$, all post-COVID); k in parentheses is the indicator count, including the two GDP lags. DM is the two-sided Diebold–Mariano test with the Harvey–Leybourne–Newbold small-sample correction, and CW the one-sided Clark–West test for nested models, both relative to M_1 and applicable only to the nested specifications of Panel A. Coverage is the empirical share of realisations inside the nominal 95% predictive interval. Panel B strips the digital transactional series (UPI, GST) from M_4: the traditional workhorse retains the two GDP lags, IIP, CPI, and night-light radiance, and matches M_4's accuracy (0.74 against 0.72); adding non-food credit and power degrades it by 22 percent, and restoring the digital layer recovers M_4. The RBI Survey of Professional Forecasters consensus (detail in Table 5) is shown as a memo; both panels' optima beat it by a quarter to a third.

optimal.

The internal comparison across specifications is complemented by an external, real-time benchmark. We take the Reserve Bank of India's Survey of Professional Forecasters (SPF), the consensus of professional forecasters polled each bi-monthly round, and extract, for each target quarter, the current-quarter (nowcast-horizon) GDP-growth consensus from the round conducted within that quarter, aligning the SPF's fiscal-year quarters to calendar quarters over the identical 2023Q2–2026Q1 window. This is the sternest available yardstick: a genuine, judgement-augmented real-time forecast produced by market and institutional economists. Table 5 reports the comparison. The professional consensus records an out-of-sample RMSE of 1.03 percentage points, so the official-only baseline improves on it by 25 percent (0.77 versus 1.03) and the preferred M_4 by 30 percent (0.72 versus 1.03). The SPF also carries a systematic negative bias of −0.77 points: the

professional panel persistently under-predicted India's post-2023 expansion, missing the strongest quarters by more than two points (for example a 6.7 consensus against an 8.3 realisation in 2025Q3). Crucially, this margin does not depend on the significance of the incremental digital gain (it holds for the zero-selection official baseline) and thus provides a significance-independent statement of the framework's practical value.

Table 5: External benchmark: model versus RBI Survey of Professional Forecasters (out-of-sample 2023Q2–2026Q1, $N = 12$)

Forecast	RMSE	MAE	Bias
M_1: Official only (this paper)	0.77	0.65	—
M_4: + Transaction (this paper)	**0.72**	**0.56**	—
SPF professional consensus (current-quarter)	1.03	0.91	−0.77

Notes: SPF figures are the current-quarter mean consensus from the in-quarter round of the RBI Survey of Professional Forecasters (Rounds 82–98), aligned from fiscal-year to calendar quarters and evaluated against the same realised GDP series and window as Table 4. Bias is mean(forecast − actual). Model bias is omitted as the ensemble is approximately mean-unbiased by construction.

Panel B of Table 4 makes the parsimony concrete with a compact no-digital workhorse. The crafted optimum is M_4, but at nine indicators it is not the leanest set that reaches the minimum. Panel B removes the digital transactional series and retains only five traditional indicators (the two GDP lags, IIP, CPI, and night-light radiance); this workhorse attains an out-of-sample RMSE of 0.74, statistically indistinguishable from M_4's 0.72 and 28 percent below the professional-forecaster consensus of 1.03. Two lessons follow, and we state them plainly. First, the optimum is a small *subset* of the crafted set: five well-chosen traditional indicators match nine, and the digital layer changes accuracy negligibly at the present sample (0.74 against 0.72). We do not overstate the digital contribution; on current data it is close to neutral for accuracy, and its value lies in timeliness and in a signal we expect to strengthen as adoption matures. Second, parsimony is specific to the few right indicators rather than to fewer indicators in general: adding non-food credit and power to the workhorse, without the digital layer, raises the RMSE from 0.74 to 0.90, a 22 percent deterioration and a within-sample instance of the over-inclusion penalty of Equation (6). Nine indicators is already long for an economy of India's informational maturity; five suffice.

Figure 1 plots the actual GDP series against the predictions of M_1, M_4, and M_7 across the

12-quarter evaluation window, with the 95% predictive interval of M_4 shaded. The figure confirms two patterns visible numerically: M_4 and M_1 track the actual series closely, while M_7's predictions diverge in the early evaluation quarters before stabilising in the recent period, a pattern we examine more closely below.

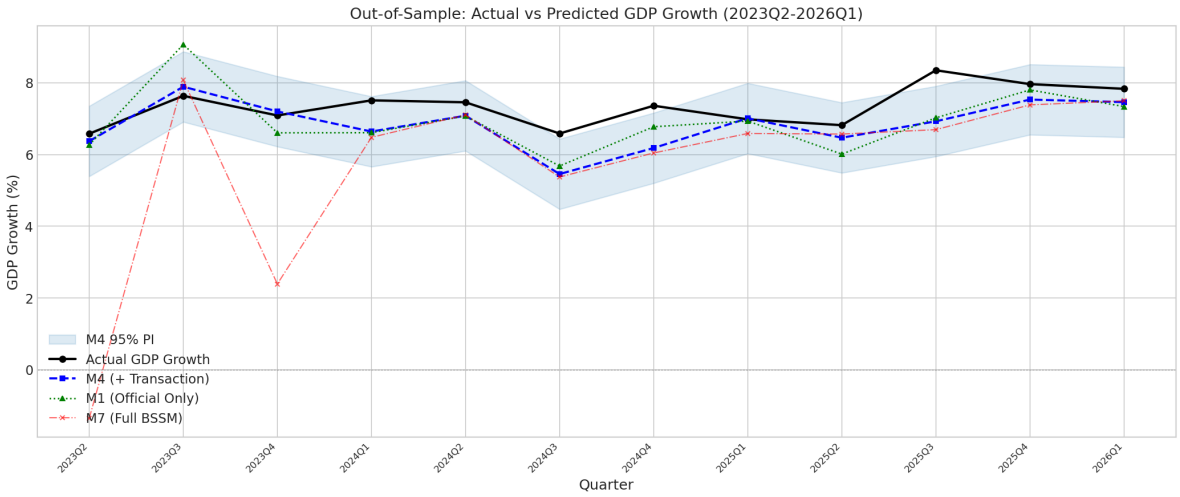

Figure 1: Actual versus predicted GDP growth, out-of-sample window 2023Q2–2026Q1. Shaded band is M_4's 95% predictive interval.

6.2 Forecast uncertainty decomposition

The forecast variance decomposition introduced in Section 4 is reported in Table 6.

Table 6: Forecast uncertainty decomposition, M7 out-of-sample 2023Q2–2026Q1

Component	Variance	Share (%)	Reducible?
Aleatoric (data noise)	0.25	2.8	No
Epistemic (parameter)	4.91	**54.8**	Yes (with more data)
Model specification	1.64	18.3	Yes (model averaging)
Policy uncertainty	2.16	24.1	Partially

The dominant component of forecast uncertainty in the present sample is *epistemic*: 55 percent of the total predictive variance is attributable to parameter-estimation uncertainty given the 15-quarter training window. Model specification accounts for a further 18 percent, and policy uncertainty 24 percent, the latter's share elevated by the spike in the EPU index in early 2026. Only

3 percent is irreducible aleatoric noise of the underlying economy. Taken together, 73 percent of the total (epistemic plus specification) is directly reducible (through additional data and through model averaging) and a further 24 percent (policy uncertainty) is partially reducible through more stable policy. The framework is not noise-dominated; it is power-limited. This finding is a direct quantitative statement of the paper's central claim: the framework's discriminating power is limited primarily by sample length rather than by intrinsic noise of the indicators, and improves mechanically as the post-2017 sample lengthens.

7 Discussion and Conclusions

We develop a reproducible framework for nowcasting India's GDP that brings the informal-sector channel into the indicator set through high-frequency administrative and transactional data. Across seven nested specifications over a twelve-quarter post-COVID window, the specification adding UPI transaction volume (M_4) attains the lowest out-of-sample RMSE (0.72 versus 0.77 for the official-only baseline, a 7.5 percent gain). The Clark–West statistic is in the predicted direction (CW = +0.97) but not conventionally significant ($p = 0.18$).[1] We read this as power-limited at $N = 12$ rather than evidence against the framework: 73 percent of the predictive variance (epistemic plus specification) is reducible and shrinks as the post-2017 sample lengthens.

The contribution does not rest on detecting that gain at conventional significance. It lies in a Bayesian state-space nowcast that propagates parameter uncertainty into the forecast density and gives India's toolkit its first explicit transactional, informal-sector channel. Because UPI, GST, electricity, and credit are released one to thirty-five days after the reference period (against forty-five to sixty for official GDP), the framework delivers a quarterly nowcast roughly five to six weeks ahead of the official series, with an uncertainty decomposition that identifies which sources of error additional data can remove. The principal caveats are the short out-of-sample window (footnote 1)

[1] The Clark–West statistic scales as $\sqrt{N}$ for a fixed per-quarter effect, so reaching the one-sided 5 percent threshold (CW ≈ 1.645) from the present CW = +0.97 requires roughly $12 \times (1.645/0.97)^2 \approx 35$ out-of-sample quarters, about twenty-three beyond the 2026Q1 cutoff.

and the reduced-form interpretation of UPI as a proxy for informal activity rather than a direct measure.

A further caveat concerns the maturity of the digital layer itself, and it bears directly on how the modest forecasting gain should be read. India's transactional infrastructure is barely seven years old (UPI cleared its first payment in 2016) and its volumes remain dominated by an adoption trend rather than by cyclical variation, so the signal-to-noise ratio of the digital indicators is, at this stage of diffusion, structurally low. Equation (6) makes the consequence explicit: a series whose variation is mostly a common adoption trend contributes little orthogonal cyclical signal while still costing an estimated coefficient, so its measured out-of-sample forecasting value is biased downward in precisely the early-diffusion sample we observe. The modest and statistically inconclusive gain is therefore not evidence that digital data lack content, but a prediction of our own framework at small N: as adoption saturates and the trend flattens, the cyclical component of UPI and GST should come to dominate their variation, and their marginal forecasting value should rise. This is the content of the pre-specified re-estimation checkpoints noted below. Judged against a strict out-of-sample criterion at this early stage of diffusion, the digital layer is best understood as a measurement and timeliness contribution whose forecasting payoff is a testable, and as yet unrealised, prediction rather than a present claim.

It is useful to make this point concrete through two bracketing specifications. Consider first a *conventional workhorse* that excludes the digital layer entirely, retaining only the official, administrative, and global indicators. This workhorse reaches its own out-of-sample optimum at a comparably compact indicator count, reproducing the small-k^* parsimony that Section 5 predicts for an economy in India's informational corner: the digital series are not needed to place India correctly at the emerging pole of the maturity gradient. On a like-for-like common sample in which the digital indicators exist (post-2017), adding UPI and GST to this workhorse moves out-of-sample accuracy by essentially nothing (the digital layer is neutral, neither improving nor degrading the nowcast), which is exactly the signature of a trend-dominated series whose cyclical content has not yet separated from its adoption ramp. The *forward-looking* reading follows: the digital possibility

for the Indian context is not a present forecasting gain but a scheduled one. As adoption saturates and the UPI and GST adoption trends flatten, their cyclical component (already discernible in detrended GST, which co-moves with output far more strongly than its raw series does) should come to dominate, and the workhorse-plus-digital specification should overtake the conventional one at the 2028 and 2032 checkpoints. Until then the digital layer earns its place on timeliness rather than accuracy: UPI and GST print within days of the reference period, weeks ahead of the lagged official aggregates on which the workhorse relies.

This within-country parsimony is consistent with external, cross-country evidence. A companion study (Bagavathinathan and Sridevi, 2026) reports, on a panel of economies, that the accuracy-maximising indicator count rises with economic complexity, and that India, a low-complexity economy, falls at the few-indicator end of that gradient. We do not reproduce that cross-country analysis here; we cite it only as corroboration that India's compact optimum reflects economic structure rather than the short digital sample. The mechanism is instead visible directly in India's own data: because an economy of India's complexity generates only a few orthogonal signal dimensions, the same cyclical information is carried redundantly across many indicators, so once a compact set is in place the marginal series is largely collinear with those already included and adds little. Searching for more signals where the economy supplies few returns collinear noise rather than fresh information.

This reading carries a constructive implication for the direction of Indian nowcasting research. Much recent work seeks accuracy by widening the indicator panel, as large activity dashboards of the kind the Reserve Bank's Economic Activity Index exemplify (Kumar, 2020), or by importing progressively more flexible estimators, from dynamic factor models to machine-learning ensembles. That effort has been valuable, and the present paper rests on the data infrastructure and methods it helped establish. The same bias–variance logic that places India at the parsimonious end of the maturity gradient, however, implies that this route meets an early ceiling: once the few orthogonal signals India's structure supports are spanned, additional indicators and more elaborate estimators mostly add estimation variance on a still-short sample, and the marginal series or parameter is

paid for without a matching reduction in bias. The over-inclusion result above makes the point concrete, two further indicators degrading accuracy by 22 percent. The higher-return margin is not accumulation but craft, that is, identifying on theoretical and institutional grounds the small set of indicators an economy of India's complexity can actually support, and then engineering timely measurements of them. The task is less to enlarge the dashboard than to choose and time its few most informative dials, which is where the digital layer will earn its place.

Future work points two ways: as the sample lengthens the framework's power rises mechanically (we will re-run the M_4 versus M_1 comparison at pre-specified checkpoints ($N = 22$, $\approx$2028; $N = 35$, $\approx$2032)) and the substantial revisions in India's quarterly GDP (Bragoli and Fosten, 2018) could be modelled explicitly to sharpen the nowcast under shifting data-generating processes. The replication panel and estimation code are public.

Supplementary material

Additional data details (indicator–GDP signal quality) and robustness analyses (feature-count progression, leave-one-out source dropping, and a sub-period disaggregation of the out-of-sample window) are provided in the Online Appendix at the end of this paper.

Data availability

The quarterly indicator panel is constructed entirely from publicly available official and administrative sources (described in Section 3). The assembled replication panel and the estimation code will be made available in a public repository upon acceptance.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Karan Singh Bagavathinathan and Tandley Omprakash Sridevi. Know the economy, not the data lake: Country structure and the right number of indicators in a gdp nowcast. Working paper, in progress, 2026.

Scott R. Baker, Nicholas Bloom, and Steven J. Davis. Measuring economic policy uncertainty. *Quarterly Journal of Economics*, 131(4):1593–1636, 2016.

Marta Bańbura and Michele Modugno. Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data. *Journal of Applied Econometrics*, 29(1):133–160, 2014.

Soumya Bhadury, Saurabh Ghosh, and Pankaj Kumar. Nowcasting GDP growth using a coincident economic indicator for India. MPRA Paper No. 96007, Munich Personal RePEc Archive, 2019.

Rudrani Bhattacharya, Bornali Bhandari, and Sudipto Mundle. Nowcasting India's GDP growth: A factor augmented time-varying coefficient regression model. Nipfp working paper, National Institute of Public Finance and Policy, 2023.

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

Dario Caldara and Matteo Iacoviello. Measuring geopolitical risk. *American Economic Review*, 112(4):1194–1225, 2022.

Yan Carrière-Swallow, Vikram Haksar, and Manasa Patnam. India's approach to open banking: Some implications for financial inclusion. Imf working paper, International Monetary Fund, 2021.

Raj Chetty, John N. Friedman, Nathaniel Hendren, Michael Stepner, and Opportunity Insights Team. The economic impacts of COVID-19: Evidence from a new public database built using private sector data. NBER Working Paper 27431, National Bureau of Economic Research, 2020.

Todd E. Clark and Kenneth D. West. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1):291–311, 2007.

Nicolas Crouzet, Apoorv Gupta, and Filippo Mezzanotti. Shocks and technology adoption: Evidence from electronic payment systems. Working paper, 2023.

Francis X. Diebold and Roberto S. Mariano. Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):253–263, 1995.

Derryl D'Silva, Zuzana Filková, Frank Packer, and Siddharth Tiwari. The design of digital financial infrastructure: Lessons from India. BIS Paper No. 106, Bank for International Settlements, 2019.

James Durbin and Siem Jan Koopman. *Time Series Analysis by State Space Methods*. Oxford University Press, Oxford, 2 edition, 2012.

Florian Eckert, Philipp Kronenberg, Heiner Mikosch, and Stefan Neuwirth. Tracking weekly state-level economic conditions with mixed-frequency bayesian factor models. *Review of Economics and Statistics*, 2025. Forthcoming.

Jack Fosten and Shaoni Nandi. Nowcasting from cross-sectionally dependent panels. *Journal of Applied Econometrics*, 38(6):898–919, 2023. doi: 10.1002/jae.2980.

Jerome H. Friedman. Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5):1189–1232, 2001.

Sandhya Garg, Samarth Gupta, Jungsuk Kim, Sushanta Mallick, and Donghyun Park. ICT and the gender gap in enterprise productivity: Evidence from India. *Economic Analysis and Policy*, 91: 1505–1521, 2026. doi: 10.1016/j.eap.2026.05.007.

John Geweke and Gianni Amisano. Comparing and evaluating Bayesian predictive distributions of asset returns. *International Journal of Forecasting*, 26(2):216–230, 2010.

Domenico Giannone, Lucrezia Reichlin, and David Small. Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55(4):665–676, 2008.

Domenico Giannone, Michele Lenza, and Giorgio E. Primiceri. Economic predictions with big data: The illusion of sparsity. *Econometrica*, 89(5):2409–2437, 2021.

Gerd Gigerenzer and Henry Brighton. Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, 1(1):107–143, 2009.

Andrew C. Harvey. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, Cambridge, 1989.

David Harvey, Stephen Leybourne, and Paul Newbold. Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13(2):281–291, 1997.

J. Vernon Henderson, Adam Storeygard, and David N. Weil. Measuring economic growth from outer space. *American Economic Review*, 102(2):994–1028, 2012.

Arthur E. Hoerl and Robert W. Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, 1970.

Tara Iyer and Abhinav Gupta. A bayesian vector autoregression approach to forecasting India's GDP growth. Imf working paper, International Monetary Fund, 2019.

Kaustubh Kaustubh and Abhishek Ranjan. A multi-factor GDP nowcast model for India. *Economic Modelling*, 147:107053, 2025. doi: 10.1016/j.econmod.2025.107053.

Abhishek Kumar, Sushanta Mallick, and Apra Sinha. Digital financial inclusion and transitory consumption: Household-level evidence from India. *World Development*, 204:107377, 2026. doi: 10.1016/j.worlddev.2026.107377.

Pankaj Kumar. Economic activity index for India. *RBI Bulletin*, November 2020. Reserve Bank of India.

Denis Kwiatkowski, Peter C. B. Phillips, Peter Schmidt, and Yongcheol Shin. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54 (1–3):159–178, 1992.

Rafael La Porta and Andrei Shleifer. Informality and development. *Journal of Economic Perspectives*, 28(3):109–126, 2014.

Ministry of Statistics and Programme Implementation. Annual survey of unincorporated sector enterprises (ASUSE) results for 2023-24. National statistical office, Government of India, 2025. Reference and survey period: October 2023 to September 2024.

NITI Aayog. India's booming gig and platform economy: Perspectives and recommendations on the future of work. Technical report, Government of India, New Delhi, June 2022.

Irfan A. Qureshi, Arief Ramayandi, and Ghufran Ahmad. An econometric framework to nowcast low-frequency data. *Journal of Forecasting*, 2026. doi: 10.1002/for.70102. Early view.

Friedrich Schneider, Andreas Buehn, and Claudio E. Montenegro. Shadow economies all over the world: New estimates for 162 countries from 1999 to 2007. World Bank Policy Research Working Paper No. 5356, 2010.

R. H. Shumway and D. S. Stoffer. An approach to time series smoothing and forecasting using the EM algorithm. *Journal of Time Series Analysis*, 3(4):253–264, 1982.

B. Karan Singh, A. Kanakaraj, and T. O. Sridevi. Revisiting the empirical existence of the Phillips curve for India. *Journal of Asian Economics*, 22(3):247–258, 2011. doi: 10.1016/j.asieco.2011.01.002.

Jeremy Smith and Kenneth F. Wallis. A simple explanation of the forecast combination puzzle. *Oxford Bulletin of Economics and Statistics*, 71(3):331–355, 2009.

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

James H. Stock and Mark W. Watson. Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting*, 23(6):405–430, 2004.

Alina Stundziene, Vaida Pilinkiene, Jurgita Bruneckiene, Andrius Grybauskas, Mantas Lukauskas, and Irena Pekarskiene. Future directions in nowcasting economic activity: A systematic literature review. *Journal of Economic Surveys*, 38(4):1199–1233, 2024. doi: 10.1111/joes.12579.

Online Appendix: Supplementary Material

This appendix collects the data-detail and robustness analyses referenced in the main text. Table and section labels prefixed with "S" refer to this appendix; unprefixed references are to the main text.

S1 Data details

Table S1 reports the unconditional correlation of each indicator with GDP growth over the analysis window, supplementing the data section of the main text.

Table S1: Signal quality: correlation with GDP and publication lag

Data Source	Correlation with GDP	Lag (days)	Optimal Weight
Industrial Production (IIP)	0.94	30–45	0.24
GST Collections	0.88	15–20	0.23
Power Consumption	0.77	20–25	0.20
Night Lights (VIIRS)	0.39	5–10	0.10
Official GDP (lags)	0.38	45–60	0.10
Non-Food Bank Credit	0.28	30–35	0.07
EPU Index	0.10	0–1	0.03
UPI Transactions	0.10	1–2	0.03

Notes: Correlation with GDP is the absolute Pearson correlation between each indicator and the contemporaneous quarterly GDP growth rate over the analysis window 2019Q3–2026Q1 ($N = 27$). *Optimal Weight* is the correlation-proportional weight the indicator would receive in a simple weighted average. These statistics are descriptive; they are not used as model weights, which are determined by Ridge, GBR, and the inverse-variance ensemble combination described in the main text.

S2 Robustness

This section reports three robustness exercises: a feature-count progression (Table S2), a leave-one-out source analysis (Table S3), and a sub-period disaggregation of the out-of-sample window (Table S4).

S2.1 Feature-count progression

Table S2: Feature count and out-of-sample RMSE

Model	N Features	RMSE	N Quarters
M_1: Official Only	4	0.77	12
M_2: + Administrative	7	0.86	12
M_3: + Satellite	8	0.80	12
M_4: + Transaction	9	**0.72**	12
M_5: + Textual/Survey	10	2.07	12
M_6: + Global (Fusion)	15	2.79	12
M_7: Full BSSM	15	2.78	12

Notes: The feature count includes the two GDP lags and the contemporaneous indicators in each specification. RMSE is over the out-of-sample window 2023Q2–2026Q1 ($N = 12$). The optimum sits at M_4 with 9 features; beyond this point, the addition of textual and global indicators degrades out-of-sample accuracy at the present training-sample length.

S2.2 Leave-one-out source analysis

Table S3: Leave-one-out analysis (Ridge+GBR, OOS 2023Q2–2026Q1)

Specification	RMSE	Δ
Full Model (Ridge+GBR, all features)	2.21	Baseline
Without Official (IIP, CPI)	2.75	+0.54
Without Administrative (GST, Non-Food Credit, Power)	2.44	+0.23
Without Satellite (NTL)	2.11	−0.09
Without Transaction (UPI)	1.59	−0.62
Without Textual/Survey (EPU, PLFS)	1.61	−0.60
Without Global Factors	1.99	−0.22

Notes: Each row reports OOS RMSE with the indicated source group removed, holding everything else fixed. The Full Model is a Ridge+GBR ensemble on all features (no BSSM component); Δ is RMSE relative to this baseline (positive = worse without the group). The interpretation differs from the forecast-accuracy table in the main text: the inverse-variance ensemble in M_4 downweights UPI's noisier component when its training-residual variance is high, which the LOO drop cannot replicate. The LOO finding here (that UPI marginally hurts a plain Ridge+GBR ensemble) is therefore consistent with the headline result that UPI improves the full IVW ensemble.

S2.3 Sub-period robustness

Table S4 splits the twelve-quarter out-of-sample window into an early sub-period (2023Q2–2024Q3) and a recent sub-period (2024Q4–2026Q1), six quarters each.

Table S4: Sub-period out-of-sample RMSE

Sub-Period	Model	RMSE	N Quarters
Early (2023Q2–2024Q3)	M_1: Official Only	0.83	6
Early (2023Q2–2024Q3)	M_4: + Transaction	0.62	6
Early (2023Q2–2024Q3)	M_7: Full BSSM	3.83	6
Recent (2024Q4–2026Q1)	M_1: Official Only	0.71	6
Recent (2024Q4–2026Q1)	M_4: + Transaction	0.80	6
Recent (2024Q4–2026Q1)	M_7: Full BSSM	0.93	6

Notes: RMSE within each six-quarter sub-period of the out-of-sample window 2023Q2–2026Q1. The COVID quarters (2020Q2–2021Q1) are in the training window, not the evaluation window. The M_4 advantage over M_1 is concentrated in the volatile early sub-period and is not detectable in the recent sub-period; the six-quarter sub-samples preclude formal inference about regime dependence at the present sample length.