'Making the Horizons Meet': A Heterodox Approach for Short-Term Inflation Forecasting

by Joice John, Saquib Hasan, Renjith Mohan, and Suvendu Sarkar ^

This article presents a framework that generates short-term inflation forecasts integrating three diverse procedures: (i) nowcasts (ii) machine learning and statistical methods and (iii) system of dynamic and stochastic equations, allowing nowcasts in the near-end of the horizon to converge to the benchmark forecasts, accentuated or delayed by persistence, spillovers, and macro-linkages. The framework is built on seasonally-adjusted disaggregated monthly data of 33 sub-groups/components of the CPI-Combined. It employs techniques such as full information matrix, machine learning and statistical models, Bayesian estimation, Kalman filtering and dynamic optimisation to produce point as well as density forecasts of inflation.

Introduction

By conducting monetary policy, central banks play a vital role in guiding economies towards macroeconomic stability and growth. While setting those policies, due to the lags in transmission and other nominal rigidities, monetary policy often focusses on forecasts of the macro variables as intermediate targets. In this context, consistent and reliable forecasts become vital for the conduct of monetary policy. More specifically, inflation forecasts are central to the monetary policy conducted by inflation targeting (IT) central banks. For forecasting inflation, there are diverse approaches available in the literature, from data dependent ones, like statistical, econometric and machine learning models to structural ones like dynamic stochastic general equilibrium (DSGE) models. In the shorter-end of the forecast horizon, data dependent models outperform the structural models, however they are not suited for policy analysis (Lucas, 1976). Structural models like DSGE are good at policy analysis but may not be favoured in terms of forecast accuracy, especially in the near-term (Del Negro and Schorfheide, 2013). At the near-end of the spectrum of any forecasting framework, there are observed data or nowcasts (since, in the near term, several auxiliary information are available) as the initial condition. However, as the horizon extends, precise information/data becomes scarcer. Hence, the forecasts become more dependent on structural characteristics like persistence, expectations, spillovers, and macro-linkages[1].

The above mentioned characteristics underscore the need for a forecasting framework, which identifies the path of convergence from observed data or nowcasts (in the near horizon) to a dynamic equilibrium[2]/benchmark forecast that is generated using an *atheoretical* framework. There can be multiple paths through which nowcast can converge to the benchmark forecasts (Chart 1).

Using dynamic optimisation, the proposed short-term forecasting model (STFM) identifies the path that allows convergence from nowcast to benchmark forecast, accentuated or delayed by persistence, spillovers from different inflation components and linkages from other macro variables like output gap, exchange rate and cost conditions. This framework also gives the flexibility to incorporate judgmental adjustments to this convergence process based on views from sectoral developments. Thus, the short-term forecasting model acts as a bridge navigating from nowcasts in the near-horizon to benchmark

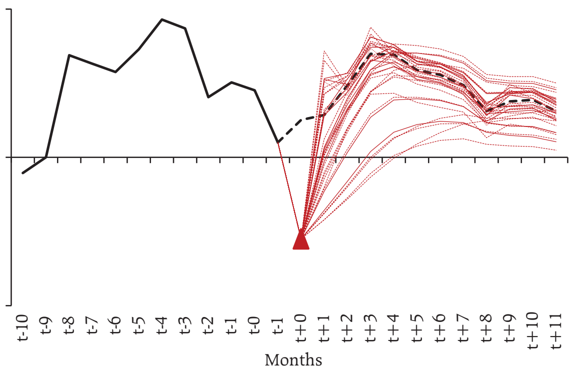
^ The authors are from the Department of Statistics and Information Management, Reserve Bank of India. The views expressed in this article are those of the authors and do not represent the views of the Reserve Bank of India.

[1] In recent times, high frequency models of macro-linkages are gaining prominence *e.g.*, Bayesian Machine Learning models and Gaussian Process and Bayesian Additive Regression Tree (BART). However, these models also, are dependent on data.

[2] Dynamic equilibrium is a state where a system is balanced, the macroscopic properties remain stable, even though changes are occurring at microscopic level.

Chart 1: From Nowcasts to Benchmark Forecast: Multiple Paths

(m-o-m, per cent)



Note: 1. This is an illustative chart and do not use any actual data, hence no axis values are mentioned.

2. Black line: Actual; Dotted Black line: Benchmark forecasts (Dynamic equilibrium); Red triangle: Nowcasts; and Dotted redlines: Possible convergence paths.

Source: Authors' Illustration.

forecasts in a short-to-medium-horizon. This hybrid architecture offers a more pragmatic and policy relevant forecasting solution. This article delves into the nitty-gritty of such a framework designed for short-term inflation forecasting in the Indian context.

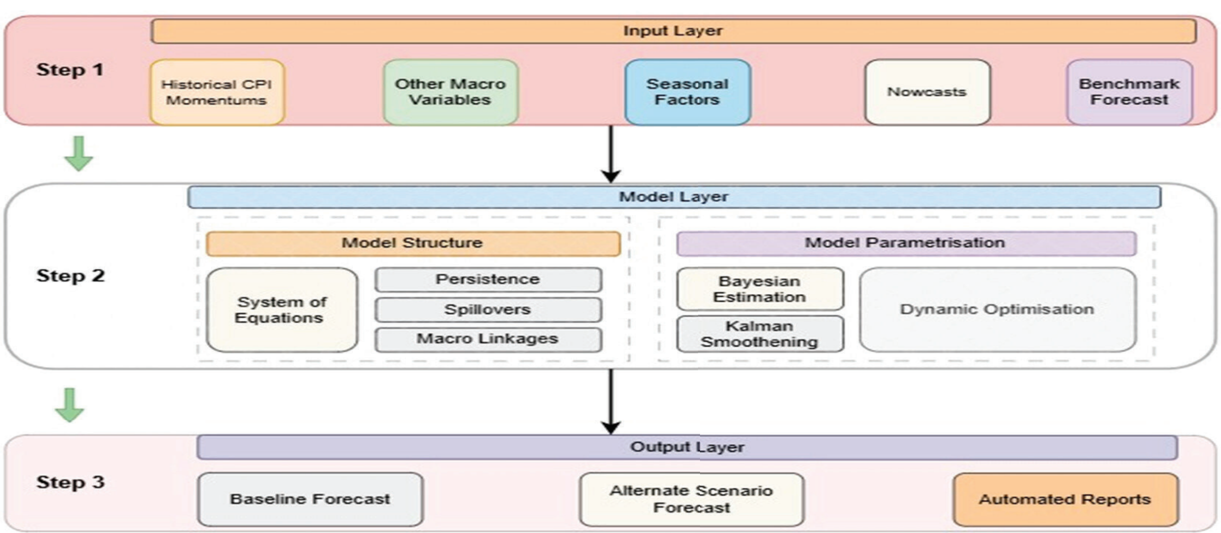
II. Short-term Inflation Forecasting Tool: An Eagle-Eye View of System Architecture and Framework Design

This section presents a high-level overview of the short-term inflation forecasting framework. It is engineered around a layered architecture that integrates diverse information sources, dynamic interlinkages and stochastic processes that allow nowcasts to converge to the benchmark forecast. The system is organized into three layers: the *Input Layer* (Step 1), the *Model Layer* (Step 2), and the *Output Layer* (Step 3) (Chart 2).

II.1. *Input Layer* (*Initial Conditions*): The *Input Layer* constitutes the historical data (CPI as well as other macro variables), seasonal factors, nowcasts and benchmark forecasts.

a) *Historical CPI Momentums*: The input layer incorporates the historical data on monthly basis for the 33 CPI sub-groups and component series[3]

Chart 2: A High-level Architecture of the Short-term Inflation Forecasting Model



Source: Authors' Illustration.

[3] Along with the officially classified 22 CPI sub-groups, 11 'fuel' items are also considered in the model framework separately. Thus, the 33 sub-groups/components are 'Cereals & products', 'Pulses & products', 'Milk & products', 'Eggs', 'Meat & Fish', 'Vegetables', 'Fruits', 'Spices', 'Oil & Fats', 'Sugar & confectionary', 'Non-Alcoholic Beverages', 'Prepared meals', 'Electricity', 'LPG', 'Kerosene-PDS', 'Kerosene-Other', 'Diesel', 'Other fuel', 'Coke', 'Firewood & chips', 'Coal', 'Charcoal', 'Dung cake', 'Housing', 'Pan, Tobacco & Intoxicants', 'Clothing', 'Footwear', 'Household', 'Health', 'Transport & communication', 'Recreation & amusement', 'Education', and 'Personal care & effects'.

for the period from February 2011 onwards. This database forms the basis for the structure and parametrisation of the model, as well as the initial condition for the short-term forecast in absence of any nowcast information.

b) *Other Macro Variables:* A set of macroeconomic variables, which includes output gap, exchange rate, commodity prices and domestic fuel costs act as the conduits of macro-linkages to headline inflation, affecting through different sub-groups/ components. These inputs are integrated within a semi-structural model to account for exchange rate passthrough, imported inflation, cost push pressures and demand-side effects on inflation.

c) *Seasonal Factors:* The model parameterisation and forecasts are carried out using the seasonally adjusted data. The seasonal adjustment process has been carried out on the month-on-month (m-o-m) changes of each of the 33 CPI series separately. These are carried out using the X-13 ARIMA-SEATS seasonal adjustment procedure[4] using the data from February 2011 onwards, with additive restrictions. Further, the average monthly seasonal factors are also computed separately for each of the 33 CPI sub-groups/components. These average seasonal factors are used in the later stage along with the seasonally adjusted m-o-m forecasts of 33 CPI sub-groups/components for generating the headline inflation forecast.

d) *Nowcasts (data dependent forecasts):* The nowcast in this forecasting framework serves as the initial condition across the 33 CPI sub-groups/ components, separately. The nowcasting process is derived from a comprehensive full information matrix constructed using all available early price signals—both quantitative (*e.g.*, daily *mandi* prices (Agmarknet, Ministry of Agriculture, Government of India (GoI)), daily wholesale and retail prices (Department of Consumer Affairs, GoI), in-house price surveys and qualitative inputs (*e.g.*, media intelligence, supply-side government measures *etc.*). This matrix acts as a real-time intelligence dashboard, capturing the most recent developments in price behaviour across these components. These nowcasts reflect the near-term momentum in prices and act as the initial conditions for the short-term forecasting model.

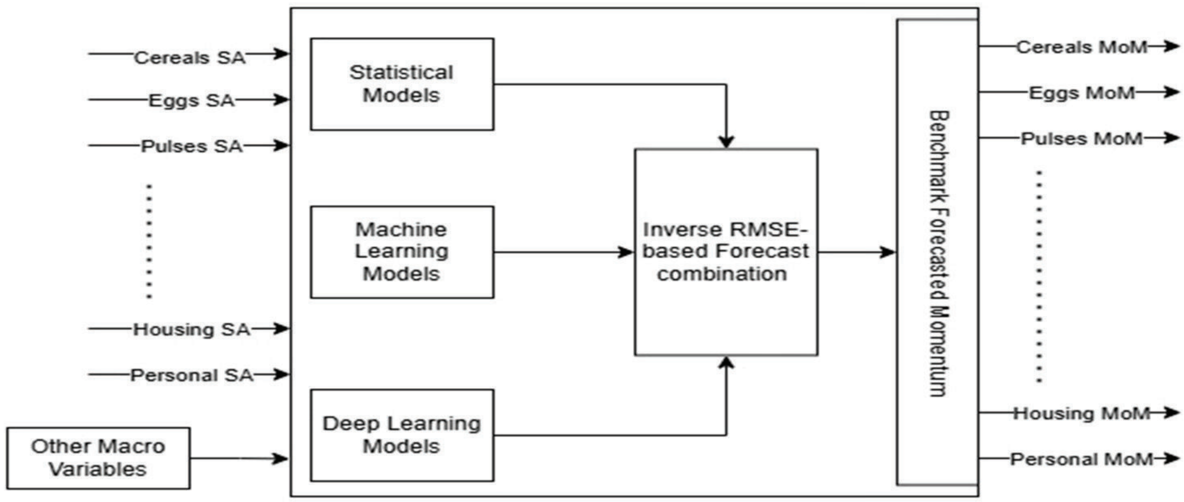
e) *Benchmark Forecasts (model dependent forecasts):* Benchmark forecasts are generated using a performance-weighted forecast combination approach (Mohan *et al.*, 2025; John *et al.*, 2020) using seasonally adjusted momentums for each of the 33 CPI sub-groups/components. For each component, this approach combines the forecasts from a *suite* of statistical, machine learning (ML) and deep learning (DL) models containing univariate and multivariate models, which includes autoregressive integrated moving average (ARIMA), vector autoregression (VAR), Bayesian VAR (BVAR), support vector machine (SVM), random forests, nonlinear autoregressive neural networks (NARNET) and long short-term memory (LSTM) models with different specifications. Thus, generating 216 forecasts for each of the 33 CPI sub-groups/components. Forecasts for individual models are then combined by weights generated using inverse of *pseudo*-out-of-sample[5] root mean squared errors (RMSEs), separately for each of these 33 sub-groups/components (Chart 3).

The benchmarks forecasts are *atheoretical* by design, making them an ideal reference point for nowcast to converge to. They reflect the information contained in historical patterns and derived out

[4] X-13ARIMA-SEATS is seasonal adjustment software produced, distributed, and maintained by the US Census Bureau (For reference material see Monsell, Lytras and Findley, 2016).

[5] In a *pseudo* out-of-sample forecasting exercise, the forecasts are generated at time *t-h* (for h = 1 to m) in the past, using only the data available till that time (*t-h*) for parametrisation of the model as well as for generating the forecast of the exogenous and endogenous variables.

Chart 3: Benchmark Forecasting Model



Source: Authors' Illustration.

of model-imposed dynamics. This makes them particularly important within the broader forecasting framework – they act as the dynamic equilibrium/ benchmark.

II.2. Model Layer (Short-term Forecasting Model): The *Model Layer* is the analytical engine of our short-term forecasting framework. It is a semi-structural model incorporating persistence, spillovers, and macro-linkages. It transforms the inputs into forecasts through Bayesian posterior updation, Kalman filtering and dynamic optimisation. This layer is bifurcated into two subsystems: *Model Structure*, and *Model Parametrisation and Estimation. Model structure* in sub-section (a) describes the equations in STFM, which are characterised by persistence, macro-linkages and interlinkages among the various CPI sub-groups/ components. *Model Parametrisation and Estimation* in sub-section (b) describes the parameter estimation process.

a) *Model Structure***:** Inflation dynamics is modelled with a set of transition equations specified for each of the 33 CPI sub-groups/components, allowing each category to respond to its own persistence, spillovers from other components (say passthrough effects from fuel prices and cost-push pressures), macro-linkages (say exchange rate passthrough to inflation and demand sensitivity) and idiosyncratic shocks. For each sub-group, inflation is governed by a system of identities and behavioural equations, such as identities linking the seasonally adjusted momentum and seasonal factors; closing identities for seasonal factors and benchmark forecasts; and dynamic behavioural equations capturing the evolution of seasonally adjusted momentums. The behavioural equations are specified as a function of lagged inflations (capturing intrinsic persistence), exogenous macro drivers (*e.g.*, output gap, cost-push pressures and exchange rate movements), spillover effects from other sub-groups/components, benchmark forecasts, and stochastic shocks. The set of equations are as below:

For all elements in

Var = {'Cereals & products', 'Pulses & products', 'Milk & products', 'Eggs', 'Meat & Fish', 'Vegetables', 'Fruits', 'Spices', 'Oil & Fats', 'Sugar & confectionary', 'Non-Alcoholic Beverages', 'Prepared meals', 'Electricity', 'LPG', 'Kerosene-

PDS', 'Kerosene-Other', 'Diesel', 'Other fuel', 'Coke', 'Firewood & chips', 'Coal', 'Charcoal', 'Dung cake', 'Housing', 'Pan, Tobacco & Intoxicants', 'Clothing', 'Footwear', 'Household', 'Health', 'Transport & communication', 'Recreation & amusement', 'Education', 'Personal care & effects'}

$$DL_{Var\{i\}} = DL_{Var\{i\}_{SA}} + DL_{Var\{i\}_{SF}} \quad (1)$$

$$DL_{Var\{i\}_{BM}} = ssDL_{Var\{i\}} + SHKDL_{Var\{i\}_{BM}} \quad (2)$$

$$DL_{Var\{i\}_{SF}} = SHKDL_{Var\{i\}_{SF}} \quad (3)$$

$$DL_{Var\{i\}_{SA}} = a^i * DL_{Var\{i\}_{SA}}\{-1\} + \underbrace{a_j^i * DL_{Var\{j\}_{SA}}}_{for\ all\ j \neq i}$$
$$+\left(1-\left(a^i+\sum a_j^i\right)\right) * \mathrm{DL}_{Var\{i\}_{\mathrm{BM}}} + b_i * DL_{Ex} + c_i * OG$$
$$+ SHKDL_{Var\{i\}_{SA}} \quad (4)$$

$DL_{Var\{i\}}$ is the m-o-m per cent change in the i^{th} variable in *Var.*

$DL_{Var\{i\}_{SA}}$ is the seasonally adjusted m-o-m per cent change in the i^{th} variable in *Var.*

$DL_{Var\{i\}_{SF}}$ is the seasonal factors of the m-o-m per cent change in the i^{th} variable in *Var.*

$DL_{Var\{i\}_{BM}}$ is the benchmark forecasts of the seasonally adjusted m-o-m per cent change in the i^{th} variable in *Var.*

OG and *Ex* are output gap and exchange rate, respectively.

Equation (1) represents the identity linking seasonally adjusted and unadjusted series. Equations (2) and (3) are used for closing the model structure. The benchmark forecasts ($DL_{Var\{i\}_{BM}}$) and seasonal factors ($DL_{Var\{i\}_{SF}}$) in the entire forecast horizon are provided as exogenous inputs. Equation (4) represents the behavioural equation encompassing persistence, spillovers, and macro impacts, which allows the convergence from nowcasts (initial condition) to the benchmark forecasts. The dimension of the short-term forecasting model is presented in Table 1.

Table 1: Dimension of the Short-term Forecasting Model

Number of CPI sub-groups/components	33
Number of equations	134
Number of variables	134
Number of shocks	101
Number of parameters	89
Number of measurement equations	68
Number of observed variables	68

Source: Authors' Estimates.

b) *Model Parametrisation and Estimation*: The model parameters are estimated using Bayesian techniques. The unobserved variables are filtered out using Multivariate Kalman Filter. Using the estimated posterior parameters and initial conditions, as provided by nowcasts, the h-period ahead forecasts are then generated using dynamic optimisation.

The Bayesian estimation is carried out using the Metropolis-Hastings[6]-Markov Chain Monte Carlo[7] (MH-MCMC) method. For each parameter in the model, a prior distribution is specified as lognormal distribution centred around a prior mode, which are identified using single equation econometric methods. The MH algorithm iteratively draws from the proposed distribution and accepts or rejects samples based on the posterior likelihood, eventually converging to the target posterior modes as defined by Bayes' rule. The estimation is governed by a set of convergence criteria, including tolerances on function values, subject to constraints and bounded by a maximum number of iterations. Once the posterior sampling is completed, the

[6] MH algorithm (Metropolis *et al.*, 1953; Hastings, 1970) is the most popular technique to build Markov chains (series of dependent samples) with a given invariant distribution. While Metropolis *et al.* (1953) requires that the proposed distribution be symmetric, Hastings (1970) generalises it to allow asymmetric distributions.

[7] MCMC methods generate Markov chains, which over time converges to a desired stationary distribution. This method is used for approximating complex distributions and estimating its parameters, even when theoretical closed-form solutions are unavailable. For details, refer to Brooks (1998).

posterior modes are computed from the MCMC draws and stored for subsequent use in filtering and forecast generation. An adaptive random-walk Metropolis (ARWM) posterior simulator[8] is used to draw samples from the prior distribution and uses estimated posterior modes to generates a large chain of iterations (here 5,00,000) to reach stationary posterior distributions. These distributions are used to generate 95 per cent credible intervals (CI)[9]. Further, the unobserved variables are filtered out using multivariate Kalman smoothing procedure[10]. Then, through a dynamic optimisation process[11], the estimated system guides nowcasts towards the benchmark forecasts, which provide the point forecasts for each of the components and groups. Further, density forecasts for each of the variables are generated using multivariate and time-simultaneous prediction bands[12]. Here, the forecast mean square error matrices are used to generate the forecast error standard deviations (SD), which in turn is applied on the point forecasts to obtain the density forecasts, assuming a normal distribution[13].

II.3. Output Layer: The Output Layer forms the final stage of the forecasting system, transforming the forecasted momentum (expressed in m-o-m per cent change) paths of each of the 33 CPI sub-groups/components—generated in the Model Layer—into forecasts of indices, year-on-year (y-o-y) inflations and contributions.

The momentum forecasts are applied to the one-period prior observed/estimated indices to recursively construct the forecasted indices for each component/sub-group. These are then aggregated into broader categories[14],—'Food & Beverages', 'Fuel & Light', and 'Core' (Ex-Food & Fuel)—using CPI weights. 'Fuel & Light' sub-group, provides an additional challenge due to the aggregation biases[15]. Hence, for 'Fuel & Light, an additional refinement is introduced. The weighted statistical moments (variance, skewness, and kurtosis) of the constituent fuel items are used as predictors for estimating the aggregation biases. From the forecasted indices, the y-o-y inflation and m-o-m rates for each sub-groups/components and at aggregated (groups and headline) levels are then calculated.

A toolbox has been developed in Matlab®, using the IRIS[16] and MikTex©[17] to support the model estimation, forecasting and output generation – including forecast tables and charts – compiled into a publication-ready report.

[8] In ARWM, a proposed distribution (here Normal) is updated adaptively using the full information accumulated so far. Due to its adaptive nature the ARWM algorithm is non-Markovian, however it has the right ergodic properties. ARWM, thus overcomes the issue of the choice of a proper distribution, which is vital for the convergence in the traditional MCMC algorithms (Haario *et al.*, 2001).

[9] Credible intervals are intervals generated from the posterior probability density function. It can be interpreted similar to the confidence interval in the frequentist approach. For *e.g.*, a 95% credible interval is having 95 per cent probability that the true value of the estimate would lie within that interval.

[10] Multivariate Kalman filter uses observed variables and stochastics noises over time to filter out unknown variables, using a multivariate state-space model, which applies the joint probability distributions in each time-step. This system level approach tends to be more accurate than those based on a single measurement variable and a single equation. A Kalman smoothing process uses both past and future values and tend to be even more accurate. For details, refer to Barratt and Boyd (2020).

[11] Dynamic optimisation involves the following steps: 1) steady state solutions are obtained using Newton-type algorithm, 2) dynamic solutions, which guides the disequilibria at any time to the steady state, are obtained using particle swarm optimizer (Eberhart and Kennedy, 1995). 3) The point forecasts are then generated using equation-selective simulator with Shanks acceleration (a non-linear algorithm which improves the rate of convergence).

[12] Multivariate and time-simultaneous prediction bands aim to capture possible outcomes for all variables at all horizons within any specified confidence level. This is used to forecasts confidence bands of different related time series by simultaneously considering the temporal uncertainty as well as their interlinkages across different variables. These are generated by estimating forecast mean square error matrices (Kolsrud, 2007).

[13] The framework can also be used to generate asymmetric confidence interval forecasts using a split-normal distribution.

[14] The 33 component/sub-group level indices are aggregated into three respective groups (Food, Fuel and Core) using the CPI-C weights. Subsequently, headline index is calculated using the group-wise CPI-C weights.

[15] The weighted vertical aggregation of the item-level indices does not match with the published 'Fuel & Light' index (Das and George, 2023).

[16] IRIS is an open-source toolbox for macroeconomic modelling and forecasting in Matlab®, originally developed by the 'IRIS Solutions Team' and currently maintained and supported by the 'Global Projection Model Network'. https://iris.igpmn.org/

[17] MikTex© is an open-source TeX /LaTeX editor for creating, typesetting, and previewing documents.

III. Results

The estimated parameters and 95 per cent CI are provided in Table 2.

Table 2: Bayesian Estimated Parameters

S. No	CPI Sub-Group	Para-meters	Prior Mode	Posterior Mode	95% CI	
					Lower	Upper
1	Cereals	a_1^1	0.53	0.70	0.47	0.95
		c_1	0.00	0.01	0.00	0.58
		a_1^{17}	0.00	0.21	0.01	0.51
2	Pulses	a_2^2	0.65	0.76	0.31	0.95
		c_2	0.10	0.02	0.01	0.84
		a_2^{17}	0.00	0.15	0.01	0.56
3	Meat & Fish	a_3^3	0.43	0.78	0.32	0.95
		a_3^{17}	0.01	0.05	0.00	0.10
4	Egg	a_4^4	0.03	0.80	0.56	0.97
		c_4	0.21	0.00	0.00	0.81
5	Milk	a_5^5	0.00	0.60	0.06	0.83
		c_5	0.17	0.00	0.00	0.63
		a_5^1	0.00	0.36	0.19	0.95
6	Oil & Fats	a_6^6	0.22	0.85	0.63	0.97
		b_6	0.34	0.01	0.00	0.88
7	Fruits	a_7^7	0.00	0.82	0.48	0.98
		a_7^{17}	0.73	0.02	0.00	0.05
8	Vegetables	a_8^8	0.00	0.53	0.29	0.76
9	Spices	a_9^9	0.67	0.85	0.48	0.97
		a_9^{17}	0.05	0.04	0.00	0.10
10	Sugar	a_{10}^{10}	0.36	0.66	0.16	0.83
		c_{10}	0.01	0.02	0.01	0.68
		a_{10}^{17}	0.01	0.25	0.02	0.64
11	Beverages	a_{11}^{11}	0.67	0.69	0.21	0.85
		a_{11}^{17}	0.00	0.01	0.00	0.10
		a_{11}^{13}	0.05	0.06	0.00	0.10
		a_{11}^{10}	0.00	0.18	0.05	0.70
12	Prepared meals	a_{12}^{12}	0.34	0.36	0.00	0.50
		c_{12}	0.02	0.01	0.00	0.77
		a_{12}^{1}	0.00	0.11	0.01	0.55
		a_{12}^{5}	0.00	0.12	0.00	0.40
		a_{12}^{15}	0.08	0.06	0.01	0.43
		a_{12}^{17}	0.00	0.03	0.00	0.28
		a_{12}^{21}	0.17	0.20	0.00	0.36
		a_{12}^{10}	0.00	0.11	0.00	0.55
13	Electricity	a_{13}^{13}	0.08	0.33	0.04	0.75
		a_{13}^{17}	0.02	0.13	0.01	0.43
		a_{13}^{19}	0.09	0.19	0.01	0.78
		a_{13}^{20}	0.00	0.31	0.00	0.48
14	LPG	a_{14}^{14}	0.32	0.74	0.48	0.95
		b_{14}	0.05	0.01	0.00	0.74
15	Kerosene-PDS	a_{15}^{15}	0.01	0.42	0.18	0.64
16	Kerosene-Other	a_{16}^{16}	0.19	0.81	0.53	0.98
		b_{16}	0.05	0.05	0.01	0.93
17	Diesel	a_{17}^{17}	0.19	0.72	0.42	0.92
		b_{17}	0.10	0.12	0.02	0.90
18	Other Fuel	a_{18}^{18}	0.00	0.78	0.35	0.92

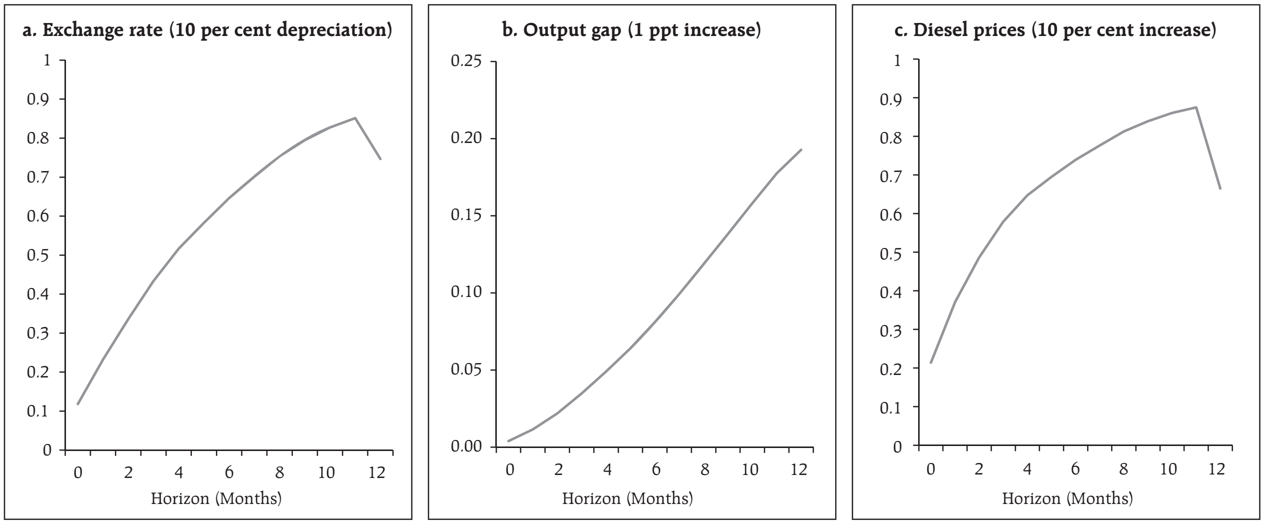
Sr. No	CPI Sub-Group	Para-meters	Prior Mode	Posterior Mode	95% CI	
					Lower	Upper
19	Coke	a_{19}^{19}	0.00	0.71	0.46	0.95
		b_{19}	0.00	0.00	0.01	0.82
20	Coal	a_{20}^{20}	0.12	0.76	0.51	0.92
		b_{20}	0.53	0.05	0.01	0.89
21	Firewood	a_{21}^{21}	0.44	0.35	0.07	0.76
		a_{21}^{14}	0.35	0.23	0.02	0.59
		a_{21}^{19}	0.00	0.32	0.01	0.59
		a_{21}^{15}	0.00	0.07	0.00	0.31
22	Charcoal	a_{22}^{22}	0.06	0.77	0.52	0.95
23	Dung cake	a_{23}^{23}	0.15	0.70	0.46	0.95
		a_{23}^{17}	0.03	0.07	0.00	0.10
		a_{23}^{15}	0.09	0.09	0.00	0.10
24	Housing	a_{24}^{24}	0.68	0.84	0.54	0.96
		c_{24}	0.01	0.00	0.00	0.74
		a_{24}^{13}	0.01	0.04	0.00	0.10
		a_{24}^{17}	0.00	0.01	0.00	0.05
25	Pan, Tobacco & Intoxicants	a_{25}^{25}	0.06	0.83	0.59	0.98
		c_{25}	0.02	0.01	0.00	0.62
26	Clothing	a_{26}^{26}	0.41	0.74	0.36	0.96
		c_{26}	0.01	0.02	0.01	0.79
		d_{26}	0.00	0.17	0.01	0.47
27	Footwear	a_{27}^{27}	0.29	0.83	0.58	0.97
		c_{27}	0.01	0.00	0.01	0.67
		a_{27}^{17}	0.00	0.01	0.00	0.05
28	Household Goods & Services	a_{28}^{28}	0.00	0.83	0.57	0.97
		c_{28}	0.02	0.01	0.00	0.65
		b_{28}	0.03	0.03	0.01	0.87
29	Health	a_{29}^{29}	0.00	0.79	0.48	0.96
		c_{29}	0.01	0.00	0.00	0.70
		a_{29}^{13}	0.00	0.05	0.00	0.10
30	Transport & Communications	a_{30}^{30}	0.28	0.78	0.46	0.97
		b_{30}	0.06	0.06	0.00	0.88
		a_{30}^{14}	0.00	0.01	0.00	0.05
		a_{30}^{17}	0.06	0.07	0.01	0.07
31	Recreation & Amusement	a_{31}^{31}	0.00	0.82	0.49	0.97
		c_{31}	0.01	0.00	0.00	0.54
32	Education	a_{32}^{32}	0.00	0.76	0.47	0.95
		c_{32}	0.03	0.01	0.00	0.66
		a_{32}^{13}	0.05	0.09	0.00	0.10
33	Personal care & effects	a_{33}^{33}	0.26	0.74	0.55	0.97
		c_{33}	0.29	0.31	0.02	0.94
		a_{33}^{6}	0.04	0.15	0.01	0.20

Note: a_j^i measures impact of i^{th} CPI sub-group on j^{th} sub-group momentum, for $i=j$, the parameter measures the persistence of j^{th} CPI subgroup, b_j measures the exchange rate pass through on the j^{th} sub-group momentum, c_j measures the impact of output gap on the j^{th} sub-group momentum

Source: Authors' Estimates.

Chart 4: Impact of Some Macro Variables on Headline Inflation (y-o-y)

(percentage points [ppt])



Source: Authors' Estimates

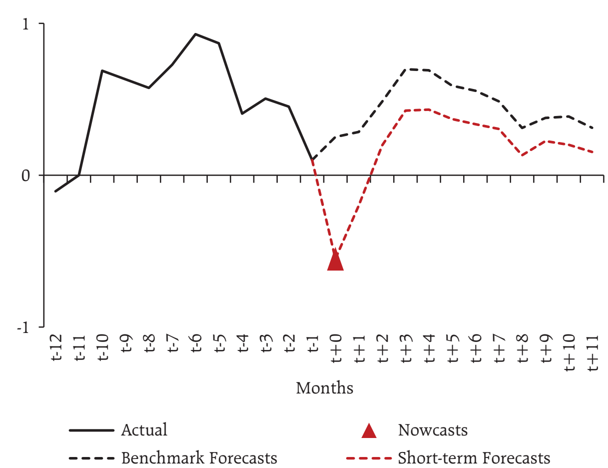
The impact of some key macro variables on headline y-o-y inflation can be derived from this framework (Chart 4). This includes both direct and indirect effects.

Ten per cent depreciation in exchange rate leads to around 80 basis points (bps) increase in inflation over a period of 12 months (Chart 4.a). If demand conditions increase by 1 per centage point (ppt), headline inflation increases by close to 20 bps over a period of one year (Chart 4.b). The ten per cent increase in diesel (pump) prices, will lead to an increase in the headline inflation by around 90 bps over a 12-month horizon (Chart 4.c).

Using one of the components as an example, Chart 5 illustrates the convergence of nowcast at period (*t+0*) to the benchmark forecast at horizon (*t+11*), generated from the short-term forecasting framework. This chart indicates that the nowcast (red triangle marker, Chart 5) of this component estimated from the high-frequency data is largely different from that emerged out of the benchmark forecast at (*t+0*) (dotted black line, Chart 5). Using nowcast as the initial condition, the short-term forecasting model allow nowcast to converge to the benchmark forecast (red dotted line, Chart 6). Here, it could be interpreted that in case of this sub-component, the difference in the near-term is largely transitory. However, the persistence, spillovers, and macro-linkages has slowed down the convergence of momentum forecast to benchmark

Chart 5: Forecast

(m-o-m, per cent)

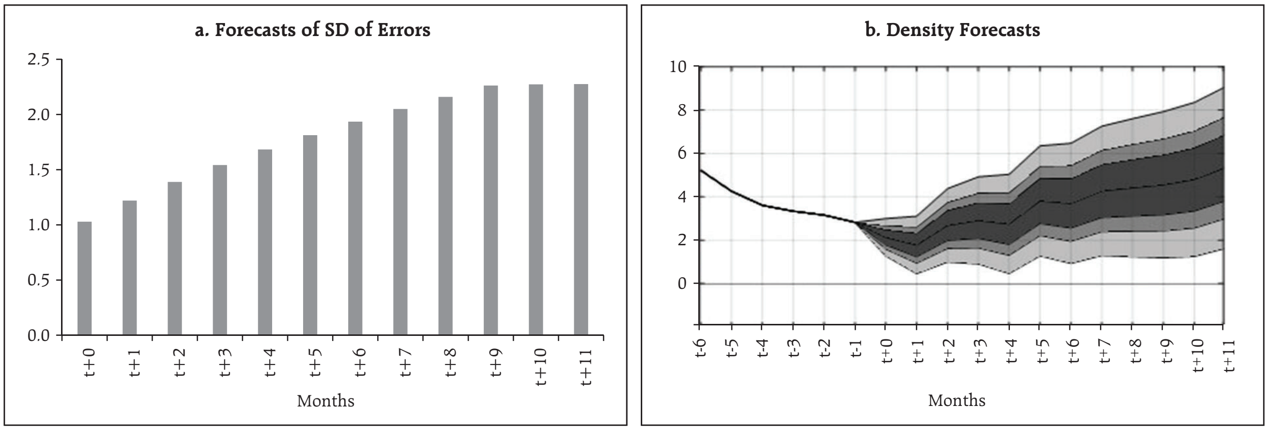


Notes: This is an illustration of the forecasted path of one of the components generated from STFM, using actual data.

Source: Authors' Estimates.

Chart 6: Density Forecasts

(y-o-y, per cent)



Notes: 1. This is an illustration of forecast path of one of the components generated from STFM, using actual data.

2. In Chart b, the thick black shaded area represents 50 per cent confidence interval, implying that there is 50 per cent probability that the actual outcome will be within the range given by the thick black shaded area. Likewise, for 70 per cent and 90 per cent confidence intervals, there is 70 per cent and 90 per cent probability, respectively, that the actual outcomes will be in the range represented by the respective shaded areas.

Source: Authors' Estimates.

forecast, even after 12-months, thus, leaving some lasting impact.

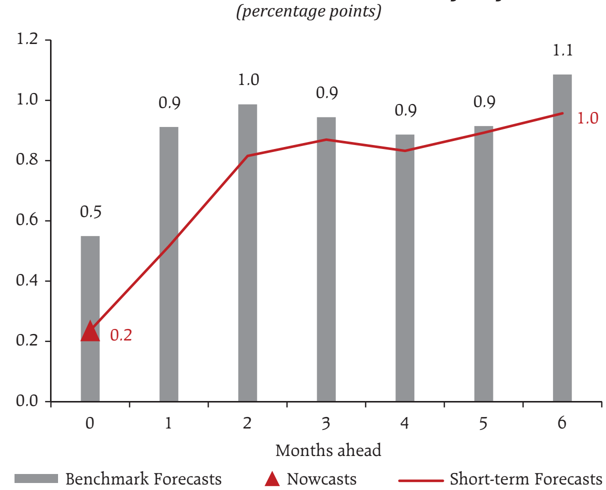
The SDs of the forecast errors for each sub-group/component that are also generated from this framework, are then used for creating the density forecasts. An illustrative example of a group in CPI basket is demonstrated in Chart 6. Chart 6.a provides the uncertainty around the point forecasts through horizons, which are measured using SDs. Estimated SDs are then used to generate the density forecast (Chart 6.b).

Finally, the evaluation of the short-term forecasting model is carried out by generating *pseudo*-out of sample root mean square errors (RMSE) (Chart 7).

Pseudo-out of sample RMSE for y-o-y headline inflation is markedly lower in the near-term compared to benchmark forecasts indicating the advantage of full information based nowcasts in the near-term. As the horizon extends the accuracy of the short-term forecasts converges to that of the benchmark forecasts. The advantage of the benchmark model (based on inflation combination approach of a large *suite* of models) in terms of accuracy for generating forecasts in short-term horizon, relative to other models is already established in the Indian context (Mohan *et al.*, 2025). Thus, the proposed framework leverages the advantage of nowcasts in the near-term, while ensuring enhanced forecast accuracy in the short-term. However, the overall accuracy of this framework depends on the accuracy of the nowcasts. This

Chart 7: *Pseudo*-out of Sample Root Mean Square Errors of Headline Inflation (y-o-y)

(percentage points)



Note: Lower RMSEs indicate better forecast accuracy.

Source: Authors' Estimates

underscores the need for a consistent and accurate framework for generating nowcasts, rather than the full information matrix–based system presented in this article. Ideally, such a framework should integrate high-frequency, spatial, and multi-source data sets–an area identified for future research.

IV. Conclusion

This paper presents a framework for short-term inflation forecasting that bridges data-driven modelling, machine learning techniques, structural hysteresis, macro-linkages, and inter-sectoral spillovers. By integrating nowcasts, benchmark forecasts, seasonal factors, and judgmental adjustments into a dynamic system of disaggregated component/sub-group level equations, this framework offers a forward-looking and granular view of inflation dynamics. The design's flexibility also enables scenario analysis. Importantly, the disaggregated architecture allows for clear attribution to inflation formation. In this framework, the enhanced forecast performance in the near-horizon stemming from nowcasts is accounted for, still preserving the advantage of statistical and machine learning models in short-horizon. It is also equipped with generating density forecasts. As such, this forecasting framework provides a powerful, yet pragmatic solution for generating short-term inflation forecasts, in an increasingly complex and uncertain environment, which are peculiar characteristics of an emerging market economy.

References:

Barratt, S. T., & Boyd, S. P. (2020, July). Fitting a Kalman smoother to data. In *2020 American Control Conference (ACC)*, 1526-1531. IEEE.

Brooks, S. (1998). Markov chain Monte Carlo method and its application. *Journal of the royal statistical society: series D (the Statistician)*, *47*(1), 69-100.

Das, P. & George, A. (2023). Consumer Price Index: The Aggregation Method Matters. RBI Bulletin, March.

Del Negro, M., & Schorfheide, F. (2013). DSGE model-based forecasting. In *Handbook of economic forecasting* (Vol. 2), 57-140. Elsevier.

Eberhart, R., & Kennedy, J. (1995, October). A new optimizer using particle swarm theory. In *MHS'95. Proceedings of the sixth international symposium on micro machine and human science*, 39-43. IEEE.

Haario, H., Saksman, E., & Tamminen, J. (2001). An Adaptive Metropolis Algorithm. *Bernoulli*, 223-242.

Hastings, W. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, *57*(1), 97-109.

John, J., Singh, S., & Kapur, M. (2020). Inflation Forecast Combinations: The Indian Experience. Reserve Bank of India Working Paper Series No. 11.

Kolsrud, D. (2007). Time-simultaneous prediction band for a time series. *Journal of Forecasting*, *26*(3), 171-188.

Lucas Jr, R. E. (1976, January). Econometric policy evaluation: A critique. In *Carnegie-Rochester conference series on public policy* (Vol. 1), 19-46. North-Holland.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics*, *21*(6), 1087-1092.

Mohan, R., Hasan, S., Roy, S., Sarkar, S., and John, J. (2025) Predicting CPI inflation in India: Combining Forecasts from a '*Suite*' of Statistical and Machine Learning Models, *RBI Bulletin*, June.

Monsell, B. C., Lytras, D., & Findley, D. F. (2016). Getting Started with X-13 ARIMA-SEATS Input Files. *US Census Bureau. Center for Statistical Research and Methodology, ML.*