NIPFP Working Paper Series

Forecasting Consumer Price Index Inflation in India: Vector Error Correction Mechanism Vs. Dynamic Factor Model Approach for Non-Stationary Time Series

No. 323
14-October-2020

Rudrani Bhattacharya, and Mriganskshi Kapoor

National Institute of Public Finance and Policy
New Delhi

Forecasting Consumer Price Index Inflation in India: Vector Error Correction Mechanism Vs. Dynamic Factor Model Approach for Non-Stationary Time Series∗

Rudrani Bhattacharya†   Mriganksha Kapoor‡

Abstract

Short to medium term forecasting of inflation rate is important for economic decision making by economic agents and timely implementation of monetary policy. In this study, we develop two alternative forecasting models for Year-on-Year (YOY) inflation in Consumer Price Index (CPI) in India using a large number of macroeconomic indicators. The YOY CPI inflation and its predictive indicators are found to be non-stationary and cointegrated. To address this issue, we employ Vector Error Correction Model (VECM) and Dynamic Factor Model (DFM) modified for non-stationary time series to forecast CPI inflation. We find that in terms of Root Mean Square Error (RMSE), the VECM model performs marginally better than the DFM model. However, both models are found to have the same predictive accuracy using Diebold-Mariano test.

JEL Classification: C32, C53.

Keywords: CPI Inflation, India, Forecasting, Vector Error Correction Model, Dynamic Factor Model.

∗This paper is written under the sponsorship of National Institute of Public Finance and Policy, New Delhi, India. We are grateful to Sudipto Mundle, Bornali Bhandari and Ajaya K. Sahu for critical and valuable comments.

†Assistant Professor, National Institute of Public Finance and Policy, New Delhi.
(Email: rudrani.bhattacharya@nipfp.org.in)
‡Birla Institute of Technology and Science, Pilani.
(Email: mrigankshikapoor@gmail.com)

1 Introduction

Accurate and timely estimation of inflation dynamics in the short as well as medium to long term in future is essential for economic agents to make decisions about consumption, savings, investment and production in real time.

Having knowledge of an accurate measure of the future trend in inflation is also essential for the central banks to implement monetary policy timely in order to maintain a moderate and stable rate of inflation. To this end, this paper discusses and compares various method of forecasting Consumer Price Index (CPI) inflation in India.

Central Statistical Organisation under the Ministry of Statistics and Programme Implementation, Government of India publishes aggregate CPI numbers and its components for base year 2011-12 for rural, urban and all India combined on a monthly basis. However these CPI series are published with a lag of one month. Hence for gauging the rate of inflation in timely manner, finding an appropriate model producing reliable inflation forecasts is of priority.

In India, monthly Year-on-Year (YOY) inflation numbers are considered for private decision making and monetary policy implementation. The YOY inflation series in India are found to be non-stationary, calling for a careful choice of techniques producing accurate, reliable as well as stable estimation of future trend of inflation in India.

To this end, we compare two multivariate techniques applicable for modeling non-stationary time series, namely Vector Error Correction Mechanism (VECM) and Dynamic Factor Model (DFM) modified for non-stationary data following Fernández-Macho (1997). We forecast CPI inflation on a monthly YOY basis for India using a large set of selected macroeconomic indicators of monthly frequency for the period April 2012 to March 2020. The indicators used in this process are YOY inflation in Wholesale Price Index (WPI), inflation for food and fuel, crude oil inflation converted into rupee terms, CPI food and core inflation, non-food credit growth, interest rate, and broad money growth.

The out of sample performance based on Root Mean Square Error (RMSE) shows that the VECM model performs marginally better than the DFM model. However using Diebold and Mariano (1995) test for comparing predictive accuracy of VECM versus DFM for non-stationary series, we find that both the models perform equally well. Both VECM and DFM for non-stationary series out perform the naive Random Walk (RW) model in terms of out of sample RMSE. Also, forecasts from VECM and DFM are found to have more predictive accuracy than the forecast from RW model at 10% and 5% level respectively.

There exists an extensive literature on forecasting inflation in India using various methods such as Vector Auto Regression (VAR) and Baysian VAR models (Biswas et al., 2010); estimation of expectation augmented Phillips curve (Kapur, 2013); and non-linear machine learning techniques in more recent times (Mitra-Thakur et al., 2016; Pratap and Sengupta, 2019).

However, to best of our knowledge, none of the studies have paid conscious attention to the non-stationary property of YOY inflation rate in India. This papers attempts to fill this gap in the literature.

The rest of the paper is organised as follows. Section 2 discusses existing literature. Section 3 describes the data used in the analysis. Section 4 outlines the VECM model and Section 5 details the DFM technique for non-stationary series. Section 6 compares forecast performance of the two models. Finally Section 7 concludes the study.

2 Literature Review

There are different schools of thought regarding modeling inflation for forecasting purposes. One way to classify the models that are used for inflation forecasting is structural and non-structural. Structural models are based on economic theory and seek to identify causal relationships among the determinants of inflation. The non-structural models like Auto Regressive Moving Average (ARMA) and Vector Auto Regression (VAR) Models have little dependence on theory.

2.1 Univariate Models

Univariate models like Random Walk (RW), Auto Regression (AR), Moving Average (MA) and Auto Regressive Moving Average (ARMA) models have been used for forecasting inflation in the literature from a very long time. RW model is often used as a benchmark model for forecast evaluation exercises. It has also been noted in the literature that simple models such as the RW and AR models are difficult to beat during periods of instability.

2.2 Multivariate Models

Over the last couple of decades, classical Vector Auto Regression (VAR) models and their modifications have also been shown to give acceptably accurate forecasts (Holden, 1995). VAR is a very popular multivariate model that expresses each variable as a function of its own lagged terms and the lagged terms of other variables in the system. When inflation rate and other indicators are non-stationary and are found to be cointegrated, the model is estimated in an error-correction framework (Sekine, 2001).

As opposed to the standard VAR where the coefficients are fixed, under the Bayesian-VAR or BVAR models, the coefficients are considered as variables that follow a known distribution (Bikker, 1998; Giannone et al., 2014). This distribution is termed as a 'prior'. Biswas et al. (2010) demonstrates the use of both VAR and BVAR in inflation forecasting for

India. The indicators chosen are quarterly IIP, Narrow Money and CPI inflation. The priors are assumed to follow multivariate normal distribution with known mean and variance-covariance matrix.

The authors then proceed to estimate the model using different sets of values for mean and co-variance, simultaneously calculating out of sample forecasts and corresponding Root Mean Square Error (RMSE) values. Based on these RMSE values, they select optimal mean and covariance values. Their findings indicate that incorporating prior information through the BVAR model significantly improves the forecasting performance.

Dynamic Factor Model following Stock and Watson (2011) is another widely used multivariate forecasting technique in the literature (Camba-Mendez and Kapetanios, 2005; Bruneau et al., 2007; Auroba and Diebold, 2010). Under these models, a large number of macroeconomic indicators are assumed to be simultaneously driven by a set of unobserved factors. These factors, following a time series process, are the underlying state variables driving the observed measurement variables. These models are estimated in State-Space form where the coefficients of the model are estimated by Maximum Likelihood Method. The latent factors are then estimated using Kalman filter and smoother.

Another approach is to breaking inflation values in two parts - trend component and cyclic component. There are several ways to decompose the trend and cyclical components through various pre-introduced models such as linear trend, the Hodric-Prescott (HP) Filter or Ravn and Uhlig modification of the HP filter (Ravn and Uhlig, 2002). Stock and Watson, (2007) assume that trend follows a random walk and cycle has an ARMA representation. Under this method, instead of forecasting headline inflation, the cyclical component for the next period is forecasted and added it to the trend. VAR models with other determinants like output gap, exchange rate and money supply are used for forecasting cyclical component, amongst other multivariate models.

2.3 Structural Models: Phillips Curve

These models express inflation as a function of output gap or unemployment gap or any other proxy for demand pull. This specification can be augmented with expected future inflation that determines the current inflation, and exchange rate capturing the component of imported inflation. One of the first instances of Phillips equation being used to construct fore casting models is Stock and Watson (1999) which comes up with an index of multiple measures of economic activity to substitute the unemployment rate gap in the traditional equation for modeling US Inflation.

One comprehensive example of a Phillips curve based forecast model applied to Indian inflation is Kapur (2013). The author has used as the foundation, the 'triangle method of inflation', an augmented Phillips equation which states that inflation depends on present and lagged values of three basic determinants: inertia (in inflation), demand and supply shocks. There is also a mention of the New Keynesian Phillips Curve in which current Inflation is dependent on expectations of inflation in the next period rather than on the previous period's inflation. In this study, quarterly Year-on-Year (YOY) inflation in Wholesale Price Index (WPI) is used as an inflation measure, while Non-food manufactured products (NFMP) inflation is regarded as a measure of core inflation. Apart from the dependent variable and output gap, some of the variables used to quantify supply shocks dynamics are rainfall, YOY growth in global prices and exchange rate etc.

Several regressions are performed using some or all of these variables with different lag periods in addition to some dummy variables marking exceptional years. Based on the R^2 values the study singles out the best performing model, which shows better accuracy than Random Walk model. The same exercise is done for NFMP inflation and it is found that it has larger coefficients of demand side variables and very small coefficients for supply side variables strengthening the case for it being a core inflation measure.

2.4 Survey Based Forecasts

Households undertake future financial decisions based on their inflation expectations. These consumption and investment decisions can in turn act as a determinant of inflation. Often, expectations can be self-fulfilling. Apart from featuring in the augmented Phillips curve equation, expectation measures can be used in data-driven multivariate inflation forecast models. While using survey based forecasts has proved to improve the performance of forecast models in countries like Brazil (Carvalho and Minella, 2012), its effect has been less pronounced in the Indian context.

A comprehensive example is Shaw (2019). The theory of expectations by Batchelor (2006) states that if the distribution of the responses coincides with a specific optimal distribution, then only the households are rational in revealing their expectations. If so, these expectations can be directly used for forecasting purposes.

Shaw (2019) shows that the Household's expectations in India are not 'rational' and therefore cannot be used for inflation forecasting directly. In this scenario, the paper attempts to use an indirect measure from the households' expectations that can be utilised for forecasting purposes. Bicchal et al. (2019) analyse expectations using high frequency data such as real- time Google Trends data and finds it adequate for forecasting. Although such data are easier to procure and are rational, these are limited in terms of their reach and cannot be assumed to be representative.

3 Data

We forecast CPI inflation on a monthly YOY basis for India using a large set of selected macroeconomic indicators for India of monthly frequency for the period April 2012 to March 2020. The Wholesale Price Index (WPI) for food is a weighted average of WPI food articles and WPI manufacturing food products. These two variables are sourced from the Office of Economic Adviser (OEA), Department of Promotion for Industry and Internal Trade, Government of India (GOI). The WPI series for energy is also sourced from OEA, GOI.

The aggregate CPI and CPI for food and energy are availed from Central Statistical Organisation, Ministry of Statistics and Programme Implementation, Government of India. The CPI core series is derived as CPI aggregate prices net of CPI food and energy prices. The Brent Crude oil prices in terms of U.S. dollar/barrel are taken from the Pink Sheet, World Bank.

The economic activities in our analysis are captured by flow of non-food credit from Scheduled Commercial Banks (SCB). This series is sourced from Centre for Monitoring Indian Economy (CMIE) Economic Outlook database. The Rupee-U.S. dollar exchange rate, interbank call money rate and broad money supply (M3) are also sourced from CMIE Economic Outlook. The expectations of households on inflation rate one year ahead surveyed by RBI is taken from CMIE Economic Outlook. The dollar price of crude oil, converted to rupee price using the rupee-dollar exchange rate is used in our model. We do not include exchange rate separately in our model.

The inflation rate forecasted in our analysis is monthly YOY inflation rate in the new combined CPI series (2012 base). The YOY change in the rest of the indicators listed above, except for the interest rate is considered for forecasting CPI inflation rate.

Under the Augmented Dickey-Fuller (ADF) test for unit roots and Phillips-Perron (PP) test which corrects ADF test for serial correlation and heteroscadasticity, we cannot reject the null hypothesis of unit root for the interest rate in level and monthly YOY changes in all the rest of the variables. Again their first differences are found to be stationary under ADF and PP tests, suggesting that these indicators are I(1) (see Tables A.1 and A.2 in Appendix A).

Under the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) Unit Root test, the null hypothesis that the series is stationary around a constant is rejected for all the indicators, except for Crude oil inflation, WPI energy inflation, and CPI core inflation at 5% and 1% level of significance. The first difference of all the indicators is found to be stationary under KPSS test.

Given that ADF and PP tests suggest that the variables used in the analysis are I(1), we also test for cointegration among them using Johansen (1991) cointegration test. The trace statistics suggest existence of one co-integrating relation among the variables at 5% and 10% level of significance (see Table A.4 in Appendix A).

4 The Model

Given one cointegration relation among the YOY changes in variables and the interest rate, we estimate the model using Vector Error Correction Model (VECM) framework as follows:

$$\Delta y_t = \mu + \alpha\beta' y_{t-1} + A_1\Delta y_{t-1} + A_2\Delta y_{t-2} + u_t \qquad (1)$$

where,

$$y_t = \begin{bmatrix} \pi_t^{\text{CPI}} \\ \pi_t^{\text{CPI core}} \\ \pi_t^{\text{CPI food}} \\ E_t\pi_{t+12} \\ \pi_t^{\text{WPI food}} \\ \pi_t^{\text{WPI energy}} \\ \pi_t^{\text{Crude oil}} \\ g_t^{\text{NFood credit}} \\ i_t \\ g_t^{\text{M3}} \end{bmatrix}$$

Here y_t denotes a vector of endogenous variables that includes aggregate CPI inflation π_t^{CPI}, CPI core inflation $\pi_t^{\text{CPI core}}$, CPI food inflation $\pi_t^{\text{CPI food}}$, inflation expectation for same month one year ahead $E_t\pi_{t+12}$, WPI food inflation $\pi_t^{\text{WPI food}}$, WPI energy inflation $\pi_t^{\text{WPI Energy}}$, crude oil inflation in rupee terms $\pi_t^{\text{Crude oil}}$, growth in non-food credit $g_t^{\text{NFood credit}}$, interest rate i_t, and growth in broad money supply g_t^{M3}.

The model is estimated after standardising the target variable i.e. CPI inflation rate and all other indicators using their respective mean and standard deviation. The parameters of the model are estimated following Johansen and Juselius (1990). Here β is the long run co-integrating vector, α is the vector of adjustment parameters, A_1 and A_2 are the cumulative short run impact parameters.[1]

The estimated coefficients in $\hat{\beta}$ represents the long run relationships among the variables. The estimated long-run relation between the variables in y_t normalised with respect to aggregate CPI inflation is shown in Table 1.

[1] The optimum lag length of 2 is chosen following the AIC criteria.

Table 1: Long run impact parameters

Variables	Long run coefficients
Aggregate CPI inflation	1.000
CPI core inflation	-0.243
CPI food inflation	-0.681
Inflation expectation	0.045
WPI food inflation	-0.143
WPI energy inflation	-0.181
Crude oil inflation	0.121
Non-food credit growth	-0.241
Interest rate	-0.189
Money supply growth	-0.116

Source: Author's estimates

We find from Table 1 that WPI food and energy inflation, growth in non-food credit, CPI food and core inflation, and growth in broad money supply positively affect aggregate CPI inflation in the long run. Higher expected inflation is found to lower CPI inflation, while we also find price puzzle as higher interest rate increases CPI inflation in the long run. We also find that rupee price inflation in crude oil prices tends to reduce CPI inflation in the long run. CPI food inflation is found to have the highest impact on headline inflation given 40% share of food in the consumption basket as well as the second round effect. CPI core inflation and Non-food credit growth follow CPI food inflation in terms of the magnitude of long run impact on aggregate CPI inflation. The effects from interest rate, WPI energy inflation and WPI food inflation are ordered next in terms of the magnitude of impacts on CPI inflation.

Table 2: Speed of adjustment to deviation from long run

Variables in first difference	Adjustment parameter	Pr(>|t|)
Aggregate CPI inflation	0.170	0.144
CPI core inflation	0.852	0.000
CPI food inflation	0.336	0.024
Inflation expectation	-0.139	0.516
WPI food inflation	-0.202	0.245
WPI energy inflation	0.089	0.290
Crude oil inflation	0.304	0.171
Non-food credit growth	-0.112	0.589
Interest rate	0.200	0.156
Money supply growth	0.436	0.084

Source: Author's estimates

Table 2 reports the estimated adjustment parameters of short run dynamics for each of the variables in system of equations (1). An adjustment parameter indicates the speed at which the short run dynamics of a variable adjust in response to a deviation from the long run relationship. For the system to converge to the long run equilibrium relationship in response to a deviation from it, the sign of adjustment parameter corresponding to a positive long run coefficient should be negative and vice versa. We find that the long run relationship affects the short run dynamics of only CPI core and food inflation among the variables in the system at 5% level of significance.

Since our aim is to forecast YOY inflation in the aggregate CPI, the estimated VECM in 1 is converted into a VAR system that expresses relations among the YOY changes in prices, non-food credit, money supply and the interest rate , after taking into account the long run relationship among them.

4.1 Forecast Performance of VECM Model

In order to assess the in-sample and out-of -sample performance of the VECM, we split our sample into the train data and test data. The train data span the period of April 2012 to September 2019, while the test data ranges for the period of October 2019 to March 2020.

4.1.1 The In-Sample Forecast Performance

Figure 1: In Sample Fit of Monthly YOY Inflation in Headline CPI: June, 2012–September, 2019

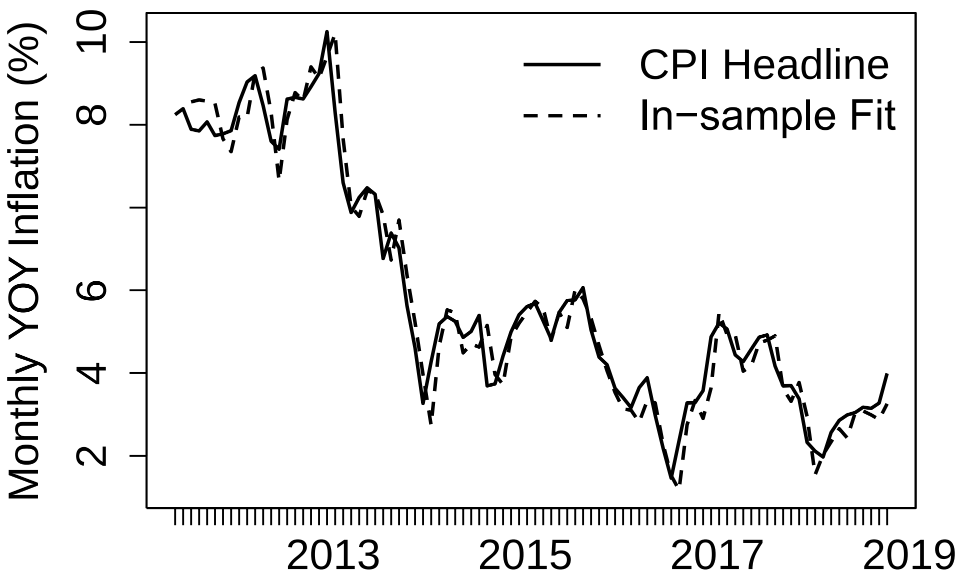

Figure 1 shows a reasonably good in-sample fit of CPI inflation from the estimated model. The in-sample Root Mean Square Error (RMSE) is found to be 0.20.

4.1.2 The Out-of-Sample Forecast Performance

In order to evaluate the out-of-sample performance of the model, we conduct one period ahead forecast by truncating the data set sequentially till February, 2020. We make a one period ahead forecast for October, 2019 using the estimated model with data for the period April, 2012 to September, 2019. Further, we make a one period ahead forecast for November, 2019 using the estimated model with data for the period, April, 2012 to October, 2019. We go on this way, until we forecast one period ahead for March, 2020 using the model estimated with data for the period April, 2012 to February, 2020. The out of sample RMSE is estimated to be 0.233.

5 An Alternative Approach: Dynamic Factor Model for Non-Stationary Time Series

In this section, we present an alternative model to forecast CPI inflation, namely, Dynamic Factor Model framework. As variables are I(1) and are conintegrated, the standard Dynamic Factor Model is not applicable here. We estimate the model using a modified version of the Dynamic Factor Model suitable for non-stationary time series following Fernández-Macho (1997). The estimated model is of the form:

The measurement equation is specified as

$$y_t = \gamma + AF_t + e_t; \qquad t = 1, 2, 3, \ldots. T \tag{2}$$

where y_t denotes a $(N \times 1)$ vector of endogenous variables that includes crude oil inflation in rupee terms $\pi_t^{\text{Crude oil}}$, WPI energy inflation $\pi_t^{\text{WPI Energy}}$, WPI food inflation $\pi_t^{\text{WPI food}}$, inflation expectation same month one year ahead $E_t\pi_{t+12}$, growth in non-food credit $g_t^{\text{NFood credit}}$, CPI food inflation $\pi_t^{\text{CPI food}}$, CPI core inflation $\pi_t^{\text{CPI core}}$, aggregate CPI inflation π_t^{CPI}, interest rate i_t, and growth in broad money supply g_t^{M3}. Here F_t is $(K \times 1)$ dimensional vector and intercept γ is a $(N \times 1)$ vector. The coefficients A is a $(N \times N)$ dimensional matrix.

The transition equation is defined as

$$F_t = \delta + F_{t-1} + u_t \tag{3}$$

where δ is $(K \times 1)$ vector of intercepts. Here e_t and u_t are modeled as Gaussian error terms $e_t \sim iid \;\; N(0, \Sigma_e)$, $u_t \sim iid \;\; N(0, \Sigma_u)$, and $E(e_t u_t) = 0$.

The DFM specification is a state-space model where the first equation, the measurement equation, describes the relation between the observed variables y_t and the unobserved state variables F_t. Equation (3) is the transition equation which describes the dynamics of unobserved state variables. In this model, unlike the conventional DFM model, the common factors are characterised by random walk with a drift. Hence, common factors are interpreted as common stochastic trend. It is assumed that $0 < K < N$. In the extreme case, where $K = 0$ or $K = N$, no common factors are present.

Table 3: Estimation results from DFM model

Dependent variable	$Ft1$ Coefficient	p-value	$Ft2$ Coefficient	p-value
CPI inflation	-22.27	0.000	-10.19	0.000
CPI core inflation	2.60	0.093	12.76163	0.000
CPI food inflation	-22.80427	0.000	-21.82104	0.000
Non-food credit growth	-11.78525	0.000	10.04526	0.000
Crude oil inflation	0.62	0.740	15.71	0.000
Inflation expectation	-9.46	0.000	-3.85	0.000
Interest rate	-13.67	0.000	6.99	0.000
Broad money growth	1.52	0.026	-5.93	0.000
WPI food inflation	-17.32	0.000	-16.02	0.000
WPI oil inflation	-0.61	0.765	17.01	0.000
	Final state	p-value	Final state	p-value
	0.024	0.016	-0.051	0.000

Source: Authors' estimates

The model estimation aims at estimating the parameters γ, A, and δ to recover the unobserved state space variable F_t. The parameters of the model are estimated using Full Information Maximum Likelihood method of Marquardt (1963). The unobserved factors are estimated using Kalman filtering technique which is a recursive algorithm that provides an optimal estimate of F_t conditional on information up to time $t-1$ and knowledge of the state space parameters γ, A, δ, Σ_e and Σ_u. The identification restrictions of A matrix is that it is formed by the first K columns of $N \times N$ lower triangular matrix and Σ_u is a diagonal matrix.

We estimate the model after standardising the variables as a ratio of deviation from respective mean and standard deviation. Hence we do not include γ and δ in our model specification. Table 3 reports the estimated coefficient and the final states.

5.1 Forecast Performance of DFM Model

In order to assess the in-sample and out-of-sample performance of the DFM, we split our sample into the train data and test data. The train data span the period of April 2012 to September 2019, while the test data ranges for the period of October 2019 to March 2020.

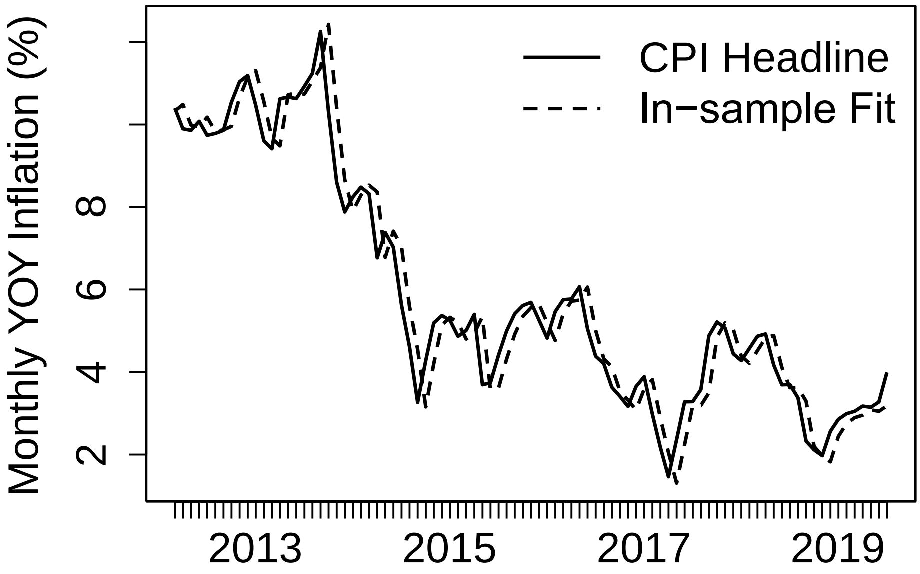

Figure 2: In Sample Fit of Monthly YOY Inflation in Headline CPI: May, 2012–September, 2019

5.1.1 The In-Sample Forecast Performance

Figure 2 shows a reasonably good in-sample fit of CPI inflation from the estimated model. The in-sample Root Mean Square Error (RMSE) is found to be 0.24.

5.1.2 The Out-of-Sample Forecast Performance

In order to evaluate the out-of-sample performance of the model, we conduct one period ahead forecast by truncating the data set sequentially till February, 2020. We make a one period ahead forecast for October, 2019 using the estimated model with data for the period April, 2012 to September, 2019. Further, we make a one period ahead forecast for November, 2019 using the estimated model with data for the period, April, 2012 to October, 2019. We go on this way, until we forecast one period ahead for March, 2020 using the model estimated with data for the period April, 2012 to February, 2020. The out of sample RMSE is estimated to be 0.355.

Table 4: Results of Diebold-Mariano Test

Model	In-sample RMSE	Out of sample RMSE	DM test H0: Two fore- casts have same predictive accuracy H1: Two fore- casts have different predictive accuracy		DM test H0: Two fore- casts have same predictive accuracy H1: Forecast 1 is more accurate than Forecast 2	
			Test Statistics	p-value	Test statistics	p-value
VECM	0.200	0.233				
DFM	0.240	0.355				
RW	0.242	0.888				
VECM vs. DFM			-0.246	0.806	-0.246	0.403
VECM vs. RW			-1.599	0.110	-1.599	0.055
DFM vs. RW			-1.730	0.084	-1.730	0.042

Source: Authors' estimates

6 Comparing Forecast Performance: VECM vs. DFM for Non-Stationary Time Series

The VECM model marginally out performs DFM model in terms of both in sample and out-of-sample RMSE (Columns 2 and 3 in Table 4). We further compare the predictive power of the two models using Diebold-Mariano test (Diebold and Mariano, 1995). We cannot reject the null hypothesis that VECM and DFM have same predictive accuracy against the alternative hypothesis that the two models have different accuracy as well as the alternative hypothesis that forecast from VECM is more accurate than forecast from DFM (Columns 4 to 7 in Table 4). We further compare forecast performance of each of VECM and DFM model with a naive Random Walk (RW) model. Forecasts from VECM and DFM are found to have more predictive accuracy than the forecast from RW model at 10% level of significance.

7 Conclusion

This study compares two alternative methods of forecasting CPI YOY inflation. To address the issue of non-stationary property of CPI YOY inflation and its predictive indicators, we employ VECM and DFM technique modified for non-stationary time series. We find that the VECM model performs marginally better than the DFM model in terms of RMSE. However both models are found to have same predictive accuracy using Diebold-Mariano test.

In the literature of inflation forecasting for India, Baysian VAR models are found to be outperforming the ordinary VAR models (Biswas et al., 2010). In the similar line and taking into account the non-stationary property of YOY inflation series in India, we propose to build a Baysian Vector Error Correction model (BVECM) following Chen and Leung (2003) and compare its predictive power with the ordinary VECM and DFM for non-stationary series as an extension of the present study.

In contrast to the statistical models, more recently with the advances in computer science, machine learning (ML) algorithms are used in forecasting studies to capture the complex and dynamic process of the macroeconomic environment (Pratap and Sengupta, 2019). Application of ML technique to forecast CPI inflation in India and comparing its predictive power with statistical models are also proposed as extensions of the present study. Finally, combining forecasts from different models are found to improve forecast performance in the literature. Exploring the combination method also belongs to our future agenda.

References

Auroba, S. B., Diebold, F. X., 2010. Real-Time Macroeconomic Monitoring: Real Activity, Inflation, and Interactions. American Economic Re- view 100 (2), 20–24.

Batchelor, R., January 2006. How Robust are Quantified Survey Data? Evidence from the United States. Cass Business School, City of London.

Bicchal, M., Durai, S., Sethu, R., 2019. Rationality of Inflation Expectations: an Interpretation of Google Trends data. Macroeconomics and Finance in Emerging Market Economies 12 (3), 229–239.

Bikker, J. A., March 1998. Inflation Forecasting for Aggregates of the EU- 7 and EU-14 with Bayesian VAR Models. Journal of Forecasting 17 (2), 147–165.

Biswas, D., Singh, S., Sinha, A., 2010. Forecasting Inflation and IIP growth: Bayesian Vector Autoregressive Model. Reserve Bank of India Occasional Papers 31 (2).

Bruneau, C., De-Bandt, O., Flageollet, A., Michaux, E., January 2007. Fore- casting Inflation using Economic Indicators: The Case of France. Journal of Forecasting 26 (1), 1–22.

Camba-Mendez, G., Kapetanios, G., 2005. Forecasting Euro Area Inflation using Dynamic Factor Measures of Underlying Inflation. Journal of Fore- casting 24, 491–503.

Carvalho, F. A., Minella, A., 2012. Survey Forecasts in Brazil: A Prismatic Assessment of Epidemiology, Performance, and Determinants. Journal of International Money and Finance 31, 1371–1391.

Chen, A.-S., Leung, M. T., May 2003. A Bayesian Vector Error Correction Model for Forecasting Exchange Rates. Computer And Operations Research 30 (6), 887–900.

Diebold, F. X., Mariano, R. S., 1995. Comparing Predictive Accuracy. Jour- nal of Business and Economic Statistics 13, 253–263.

Fern´andez-Macho, F. J., 1997. A Dynamic Factor Model for Economic Time Series. Kybernetika 33 (6), 583–606.

Giannone, D., Lenza, M., Momferatou, D., Onorante, L., July-September 2014. Short-term Inflation Projections: A Bayesian Vector Autoregressive Approach. International Journal of Forecasting 30 (3), 635–644.

Holden, K., May 1995. Vector Auto Regression Modelling and Forecasting. Journal of Forecasting 14 (3), 159–166.

Johansen, S., 1991. Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. Econometrica 59 (6), 1551–1580.

Johansen, S., Juselius, K., 1990. Maximum Likelihood Estimation and Inference on Co-integration-with Application to the Demand for Money. Oxford Bulletin of Economics and Statistics 52, 169–210.

Kapur, M., April 2013. Revisiting the Phillips curve for India and inflation forecasting. Journal of Asian Economics 25, 17–27.

URL https://doi.org/10.1016/j.asieco.2012.12.002

Marquardt, D. W., June 1963. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. Journal of the Society for Industrial and Applied Mathematics 11 (2), 431–441.

Mitra-Thakur, G. S., Bhattacharya, R., Srakar-Mondal, S., 2016. Artificial Neural Network Based Model for Forecasting of Inflation in India. Fuzzy Information and Engineering 8 (1).

Pratap, B., Sengupta, S., 2019. Macroeconomic Forecasting in India: Does Machine Learning Hold the Key to Better Forecasts? Working Paper WPS (DEPR): 04/ 2019, Reserve Bank of India.

Ravn, M. O., Uhlig, H., 2002. On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations. The Review of Economics and Statistics 84 (2), 371–380.

Sekine, T., 2001. Modelling and Forecasting Inflation in Japan. Working Paper WP/01/82, International Monetary Fund.

Shaw, P., 07 2019. Using Rational Expectations to Predict Inflation. RBI Occasional Papers 40 (1), 85–104.

Stock, J. H., Watson, M. W., 1999. Forecasting Inflation. Journal of Monetary Economics 44 (2), 293–335.

URL https://ideas.repec.org/a/eee/moneco/v44y1999i2p293-335.html

Stock, J. H., Watson, M. W., 2007. Why Has U.S. Inflation Become Harder to Forecast? Journal of Money, Credit and Banking 39 (S1), 3–33.

Stock, J. H., Watson, M. W., 2011. Dynamic Factor Models. In: Clements, M. P., Hendry, D. F. (Eds.), Oxford Handbook on Economic Forecasting.

A Appendix A

Table A.1: Augmented Dickey-Fuller Test Results

Variable	Test statistic	Variable	Test statistic
Crude oil YOY inflation	-2.74	Δ Crude oil YOY inflation	-6.20
WPI energy YOY inflation	2.59	Δ WPI energy YOY inflation	-6.18
WPI food YOY inflation	-2.32	Δ WPI food YOY inflation	-7.00
Expected inflation (one year ahead)	-1.59	Δ Expected inflation	-9.59
Non-food credit YOY growth	-1.96	Δ Non-food credit YOY growth	-11.27
CPI food YOY inflation	-2.42	Δ CPI food YOY inflation	-6.84
CPI core YOY inflation	-3.68	Δ CPI core YOY inflation	-5.02
CPI YOY inflation	-1.68	Δ CPI YOY inflation	-6.04
Call money rate	-0.85	Δ Call money rate	-6.82

Source: Author's estimates

The YOY changes in the variables and the interest rate are tested with the null hypothesis of unit root with a drift but no trend. The critical values under this null hypothesis are -3.51, -2.89 and -2.58 at 1%, 5% and 10% level of significance respectively.

The first difference of YOY changes in the variables and the interest rate are tested with the null hypothesis of unit root with no drift or trend. The critical values under this null hypothesis are -2.60, -1.95 and -1.61 at 1%, 5% and 10% level of significance respectively.

Table A.2: Phillips-Perron Test Results

Variable	Test statistic	Variable	Test statistic
Crude oil YOY inflation	-2.20	Δ Crude oil YOY inflation	-6.27
WPI energy YOY inflation	-1.98	Δ WPI energy YOY inflation	-5.34
WPI food YOY inflation	-1.92	Δ WPI food YOY inflation	-7.03
Expected inflation (one year ahead)	-1.59	Δ Expected inflation	-9.57
Non-food credit YOY growth	-1.80	Δ Non-food credit YOY growth	-11.38
CPI food YOY inflation	-1.88	Δ CPI food YOY inflation	-6.35
CPI core YOY inflation	-2.90	Δ CPI core YOY inflation	-6.38
CPI YOY inflation	-1.78	Δ CPI YOY inflation	-6.83
Call money rate	0.97	Δ Call money rate	-7.88

Source: Author's estimates

The YOY changes in the variables, the interest rate and their first difference are tested with the null hypothesis of unit root with a drift but no trend. The critical values under this null hypothesis are -3.50, -2.89 and -2.58 at 1%, 5% and 10% level of significance respectively.

Table A.3: KPSS Test Results

Variable	Test statistic	Variable	Test statistic
Crude oil YOY inflation	0.26		
WPI energy YOY inflation	0.289		
WPI food YOY inflation	0.670	Δ WPI food YOY inflation	0.078
Expected inflation (one year ahead)	1.71	Δ Expected inflation	0.063
Non-food credit YOY growth	1.02	Δ Non-food credit YOY growth	0.073
CPI food YOY inflation	1.300	Δ CPI food YOY inflation	0.114
CPI core YOY inflation	0.107		
CPI YOY inflation	1.710	Δ CPI YOY inflation	0.160
Call money rate	2.219	Δ Call money rate	0.239

Source: Author's estimates

The null hypothesis is that the series is stationary around a constant. The critical values under this null hypothesis are 0.739, 0.463 and 0.347 at 1%, 5% and 10% level of significance respectively.

Table A.4: Johansen Cointegration Test Results

Rank	Test statistic	10%	5%	1%
$r \leq 9$	2.77	6.50	8.18	11.65
$r \leq 8$	7.92	15.66	17.95	23.52
$r \leq 7$	15.19	28.71	31.52	37.22
$r \leq 6$	25.57	45.23	48.28	55.43
$r \leq 5$	43.50	66.49	70.60	78.87
$r \leq 4$	68.83	85.18	90.39	104.20
$r \leq 3$	100.00	118.99	124.25	136.06
$r \leq 2$	133.89	151.38	157.11	168.92
$r \leq 1$	181.34	186.54	192.84	204.79
$r = 0$	233.70	226.34	232.49	246.27

Source: Author's estimates

MORE IN THE SERIES

- Pandey, R., Sapre, A., and Sinha, P., (2020). What do we gain from Seasonal Adjustment of the Indian Index of Industrial Production (IIP)? W.P. No. 322 (October).
- Bhattacharya, R., Sen Gupta, A., and Sikdar, S., (2020). Building Infrastructure to Promote Inclusive Growth, W.P. No. 321 (October).
- Bose, S., Bera M., and Ghosh, P., (2020). Centre-State Spending on Elementary Education: Is it Complementary or Substitutionary? W.P. No. 320 (September).

Rudrani Bhattacharya, is Assistant Professor, NIPFP
Email: rudrani.bhattacharya@nipfp.org.in

Mrigankshi Kapoor, Birla Institute of Technology and Science, Pilani.
Email: mrigankshikapoor@gmail.com

National Institute of Public Finance and Policy,
18/2, Satsang Vihar Marg, Special Institutional Area (Near JNU),
New Delhi 110067
Tel. No. 26569303, 26569780, 26569784
Fax: 91-11-26852548
www.nipfp.org.in