Forecasting Digital Payment Systems Data for Nowcasting Macroeconomic Variables in India – A Deep Learning Perspective

Naveen Kumar Singh[1], Yash Agarwal[2], Sonali Adki[3]

Abstract

The recent worldwide development and widespread use of digital payment systems has provided an opportunity to explore new sources of data for the monitoring of macroeconomic activity. In India, digital payment systems data are available at a higher frequency (daily). In this paper, therefore, we analyse the usefulness of higher frequency data collected from different payment systems for nowcasting quarterly macroeconomic variables. We take advantage of the availability of such higher frequency data firstly, to forecast higher frequency payment system indicators, and use the forecasts to nowcast lower frequency economic variables such as Gross Domestic Product (GDP) and Private Final Consumption Expenditure (PFCE). We use new state of the art deep learning models like Prophet-with-XGBOOST-errors, Feed forward autoregressive neural networks and their weighted ensembles to forecast payment system data. We forecast payment systems like Real Time Gross Settlement (RTGS) which is typically a large value payment system, retail payment systems like Unified Payments Interface (UPI), National Electronic Funds Transfer (NEFT), Credit and Debit cards, Immediate Payment Service (IMPS), etc. through these methods. We use Autoregressive Integrated Moving Average (ARIMA) and Ordinary Least Squares (OLS) method for nowcasting GDP and PFCE.[4]

Key words: GDP, PFCE, payment systems in India, forecasting, nowcasting, machine (deep) learning, ANN, supervised learning.

JEL classification: E17, E32, E37, C53

1. Introduction

India has been enjoying a healthy evolution of payment systems over the past three decades. This has been the result of the measured road maps periodically adopted by the Reserve Bank of India, as a developer in the initial years and as a catalyst and facilitator in later years. Though the advancements in the payment systems were gradual in the early days, the last two decades of this century have truly witnessed a revolution. From barter system to Unified Payments Interface (UPI), payment systems in India have come a long way.

Ever since the global recession, interest in new macroeconomic forecasting tools, especially those based on monetary and financial information, has been increasing. On the back of the developments of computational tools for storing and elaborating large-scale data sets, analysts are focusing on the pursuit of new, timely, and reliable information in order to improve the forecasting ability in real time. This is especially important during crisis times like COVID-19 when real time payment systems data

[1] Manager, Department of Payment and Settlement Systems
[2] MBA student, Narsee Monjee Institute of Management Studies and Summer Intern at Reserve Bank of India
[3] Deputy General Manager, Department of Payment and Settlement Systems
[4] The views expressed in the paper are those of the authors and do not reflect the views of the Reserve Bank of India

can shape the policy inputs significantly for lag of reporting of macroeconomic variables like GDP, inflation, Private Final Consumption Expenditure, etc.

Forecasting of payment system transactions is also important to derive important macroeconomic insights into the future. While on the one hand, it helps in scheduling the servers of financial institutions for any unanticipated bump in payment system volume, on the other hand it is also widely used for nowcasting of macroeconomic variables like GDP and PFCE, just to name a few. Traditionally, forecasting methods like Autoregressive Integrated Moving Average (ARIMA), Trigonometric seasonality, Box-Cox transformation, ARIMA errors, Trend, Seasonal components (TBATS), Error Trend and Seasonality, or exponential smoothing (ETS) and their seasonal variants have been used for forecasting time series. However, they fall short of expectations when it comes to forecasting using daily data, which is highly volatile. Moreover, they generally provide good forecasts for only short term.

For nowcasting, the general practice is to nowcast a lower frequency variable using a higher frequency variable. However, the data for the higher frequency variable should be available for the entire period for which the lower frequency variable is to be nowcasted. Our research shows that we can nowcast lower frequency macroeconomic variables like GDP and PFCE with high degree of accuracy for near future period using forecasted values of higher frequency payment system indicators. We use Autoregressive Integrated Moving Average (ARIMA) and Ordinary Least Squares (OLS) method for nowcasting purpose. Among a set of frequently used predictors of macroeconomic activity, payment system data delivers some of the best nowcasts and one-quarter-ahead forecasts of GDP, which reinforces the potential usefulness of this novel type of data among policymakers and practitioners. We show that payment system data can help predict economic activity in India by employing deep learning methods to forecast quarterly macroeconomic variables using high-frequency data.

Table 1: Overall payment system indicator statistics for July 2022

Month	July-18	July-21	June-22	July-22	MoM %	YoY %	CAGR
Digital volume	170.02	527.12	829.76	877.58	5.76	66.49	50.73
Digital value	133.19	142.00	169.49	161.25	-4.86	13.56	4.90
Total volume	179.43	533.08	835.70	883.47	5.72	65.73	48.96
Total value	139.94	147.53	175.48	167.05	-4.81	13.23	4.53
Retail volume	178.34	531.40	833.75	881.57	5.74	65.90	49.11
Retail value	27.92	40.12	51.92	51.53	-0.75	28.46	16.55
Digital retail volume	168.92	525.44	827.81	875.69	5.78	66.66	50.89
Digital retail value	21.17	34.58	45.93	45.73	-0.42	32.25	21.23

Volume in Crore and Values in INR Lakh Crore; CAGR is for last 4 years

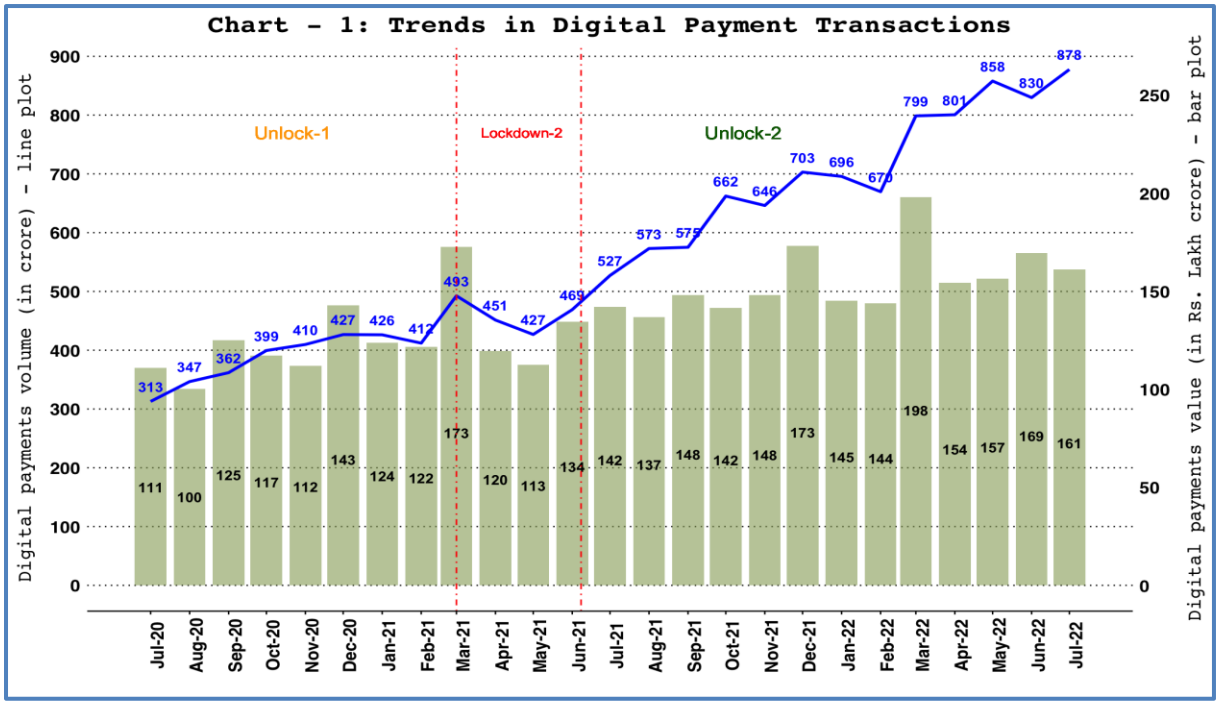
Note: : 1. Digital payments include RTGS (customer and inter-bank transactions), retail electronic clearing – NEFT, IMPS, NACH (credit, debit and APBS), card paymenttransactions (excl. cash withdrawal), PPI payment transactions(excl. cash withdrawal), UPI (including BHIM & USSD), BHIM Aadhaar Pay, AePS fund transfer and NETC (linked to bank accounts).

2. Total payments include Digital Payments plus paper clearing

3. Retail payments include NEFT, IMPS, NACH (credit, debit and APBS), card payment transactions (excl. cash withdrawal), PPI payment transactions (excl. cash withdrawal), UPI (including BHIM & USSD), BHIM Aadhaar Pay, AePS fund transfer and NETC (linked to bank accounts) and paper clearing.

4. Retail Degital include Retail payments minus paper clearing

During the last decade, the country has witnessed the introduction of innovative payment systems, the entry of non-bank players, and a gradual shift in customer behavior from cash to digital payments (Booklet on Payment Systems -January 2021, RBI). In the last four years (Feb'18 to Feb'22), the digital payment transactions volume and value increased at CAGR of 69% and 9.6% respectively. Even during the COVID-19 pandemic, digital payment transactions increased by 28.7% in volume in FY 2020-21 as compared to FY 2019-20. During the same period, the credit and debit card payments volume increased at a CAGR of 20.2% and 7.9% respectively. On the other hand, the UPI and IMPS payment volume increased at a CAGR of 66.4% and 241.6% respectively. Also, RTGS and NEFT increased by 16.7% and 28.0% respectively during the same period. Thus, even within the digital payment systems, there has been huge volatility in performance. This makes the task of forecasting payment systems particularly challenging.



Forecasting of payment system indicators data can be very crucial for making policy decisions by central banks. Moreover, the impact of policy decisions can be gauged in a short term by estimating the gap between forecasted and real figures. There is a growing requirement to nowcast macroeconomic variables like Gross Domestic Product (GDP), Private Final Consumption Expenditure (PFCE), Gross Fixed Capital Formation (GFCF), etc. These variables are typically calculated with a lag of a quarter. The Payment systems trace economic transactions, so they could be considered important indicators of economic activity. In India, payment systems data is available at monthly and daily frequency. Therefore, estimating the relationship between these macroeconomic variables and payment systems, and the availability of accurate forecasting estimates for the payment systems is the *sine qua non* for a good nowcasting model.

This research paper is an effort to shift the frontier in the forecasting payment systems data and use it to nowcast macroeconomic variables like GDP and PFCE.

Table 2: Description of Payment System Indicators in India	
Name	Description
UPI	Unified Payments Interface is an instant real-time payment system facilitating inter-bank transactions operated by National Payments Corporation of India (NPCI).
RTGS	Real Time Gross Settlement is a funds transfer systems where transfer of money takes place from one bank to another on a real time and on gross basis. This was introduced in in 2004 and settles all inter-bank payments and customer transactions above ₹Z2 lakh. From December 2020. it is available 24x7 throughout the year.
NEFT	National Electronic Funds Transfer was introduced in November 2005. It is a secure system for facilitating one-to-one funds transfer requirements of individuals / corporates. The NEFT system provides for batch settlements at

	half-hourly intervals, thus enabling near real-time transfer of funds. From December 2019, it is available 24x7 throughout the year with half-hourly settlements.
NACH- Debit	National Automated Clearing House was formed to consolidate multiple Electronic Clearing Service systems running across the country into one centralised system. NACH Debit operates to collect payments from many accounts to one destination account e.g. collection of various utility payments pertaining to telephone, electricity, water and gas charges etc and periodic instalments towards loans, investments in mutual funds, insurance premium etc.
Credit Cards	A credit card is a payment card issued to users (cardholders) to enable the cardholder to pay a merchant for goods and services based on the cardholder's accrued debt (i.e.. promise to the card issuer to pay them for the amounts plus the other agreed charges).
Debit Cards	A debit card is a plastic payment card that can be used instead of cash when making purchases. It is similar to a credit card, but unlike a credit card. the money is immediately transferred directly from the cardholder's bank account to pay for the transaction.
Prepaid Payment Instruments (PPI)	Prepaid Payment Instruments are payment instruments that facilitate purchase of goods and services, including financial services, remittance facilities. etc., against the value stored on such instruments. It may be issued as cards, wallets. and any such electronic or digital form I instrument which can be used to access the PPI and to use the amount therein.

2. Literature review

Machine learning models outperform traditional models with better performance and can identify turning points well before its competing methods which can help forecasters with more guidance of likely future fluctuations (Hall, 2018). Machine learning is quickly becoming familiar, as the high-performance algorithms are showing outstanding outcomes across many fields in both industry and academia (Soohyon Kim, Bank of Korea's WP-2020). Machine learning combines elements from computational statistics, mathematical optimization, pattern recognition, and predictive analytics (Chakraborty and Joseph, 2017), by which the algorithms train themselves through "learning" the latent pattern hidden in the data. As deep learning - it is more advanced algorithm as a branch of artificial neural networks – is being introduced, and as more-detailed data becomes available, machine learning approaches can be stretched to cover a wider variety of tasks. However, we acknowledge that machine (deep) learning is familiar within the field of economics, while even statisticians have accepted this revolutionary approach as being part of their methodologies (Athey and Imbens, 2019). For decades, economics and finance have heavily relied on econometric models for empirical analyses, but we witness nothing but somewhat meaningful progress as being made by, say, "inverting a covariance matrix (Lopez de Prado, 2018a)". While empirical analyses in economics and finance adhere to such conventional approaches, machine learning is proving itself as a possible alternative to econometrics, especially when it is difficult to obtain the answers to the questions through conventional approaches. Machine learning is a set of algorithms that we can program to train themselves to have desirable predictive powers concerning future outcomes.With its lack of a structural background based on economic theories, we can still benefit from its predictive power as long as the research objects are all about prediction or forecasting. It is widely acknowledged that

policymakers should not miss out on the opportunities provided by big data – described by some as the new oil of the 21st century [Economist (2017)]. Public institutions are not the main producers of big datasets, and some of this information may have little relevance for their daily work. Yet, central banks are increasingly dealing with "financial big data" sources that impinge on a wide range of their activities.

There are many areas in which the use of machine learning, deep learning, and other advanced techniques are used for time series forecasting. These techniques are proven in the past to give the accurate result and many authors have incorporated such methods in their research and analysis. For the sake of this research, an exhaustive literature survey has been carried out to highlight the use of advanced time series forecasting techniques in the light of daily payment systems data. Some of them are included below.

2.I. Studies in Forecasting of financial time series

ANN Models such as Multilayer Feed Forward Neural Network (MLFFNN) and Nonlinear Autoregressive models with Exogenous Input (NARX) and Time Series Econometric Models like Generalized Autoregressive Conditional Heteroskedastic (GARCH) and Exponential Generalized Autoregressive Conditional Heteroskedastic (EGARCH) used to forecast daily Indian Rupee to US Dollar Exchange Rate with independent variables from both capital and current accounts of Balance of Payments (BoP)(Chaudhuri & Ghosh, 2016). Sudhakar & Mahadevan (2019) nowcasted Indian Real GDP using ARIMA-type modelling with daily RTGS volume, demonstrating that RTGS and GDP data are linked. However, Newaz (2008) used the Box-Jenkins methodology for building ARIMA Models, Exponential smoothing, Naive 1 and Naive 2 models on sample data issued by International Financial Statistics by IMF. The comparison of forecasts was done based on Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and Mean Square Error (MSE), in which the ARIMA model performed the best.

Time series modelling for banking transactions in India can be seen in (KullayaSwamy, Sarojamma, & Kiran Kumar, 2018) where authors used monthly data of NEFT and RTGS transactions from 2010-2017 and applied a seasional ARIMA model to predict the volume and value of NEFT and RTGS till December 2020 . Similarly, , Sharma & Bhatia (2012) used monthly data of 2009-2011 for short-term transactions (NEFT) forecasting using Seasonal ARIMA model and Multiplicative Decomposition Model. The predictions of the two models compared using the MAPE in which the multiplicative model slightly outperforms the SARIMA model. On the other hand Karabiber & Xydis (2019) forecasted electricity price of Denmark-West region for 212 days of 2017 with past daily values from 2016. They used ARIMA, ANN, and TBATS algorithm and their ensembles and models evaluated based on MAE and RMSE. To predict daily data of financial transactions of ING Bank of Netherlands, Corstjens, Bakhshandeh, Kahraman, & Bosman (2019) used multiple linear regression. They combined calendar features with that of the domain of IT systems of ING Bank and created a positive impact on the bank with high customer satisfaction.

Yenidoğan, Çayir, Kozan, Dağ, & Arslan (2018) presents Bitcoin forecasting with the help of ARIMA and Prophet models. Both the models are compared based on their performance metrics. Jain & Prasad (2020) present Prophet and XGBoost Algorithms to resolve the issue of estimation of traffic in telecom

systems having time-series data. They show that these two models hold best to resolve the issue of twining of data and provides efficient data modelling to support network infrastructure. Jackson, Muhammad, & Salaser (2021) highlights the Seasonal Persistence technique to forecast the sales of Europe cashback websites. They applied models such as ARIMA, XGBoost, FbProphet on the sales data where the XGBoost algorithm stood the best with the lowest MAPE. The sales data forecated using a composite of GRU-Prophet model with an attention mechanism to capture non-linear and linear features by Li., Yang, Zhu, & Zhnag . The authors concluded that the composite model outperforms traditional ARIMA, Prophet, RNN, LSTM type models in terms of accuracy in predictions.

The ARIMA model is good only for less varied data and fails to provide good results in case of a sudden change in the data such as due to economic instability (Guha & Bandyopadhyay (2016) and Banerjee (2014)). Moreover, the ARIMA model takes linear historical data as an assumption, though, in real-world scenarios data can be non-linear over time.There is plenty of research material available on the nowcasting of macroeconomic variables. In the Indian context, large payments through electronic payment systems positively impact economic growth. For instance, Raju & Balakrishnan (Spring 2019) used ARIMA-type modelling to nowcast Indian Real GDP with monthly RTGS volume data and shows that RTGS and GDP are co-integrated and concluded that data of payments system indicators can be used to enhance GDP forecasts. The payment systems are key indicators of economic activity of a country and found that contribution of Payment systems flows in predicting GDP is non-negligible (Aprigliano, Ardizzi, & Monteforte, 2019). They nowcast the monthly data of Italy's Retail Settlement Systems and the country's GDP and its components using mixed frequency factor-based model. The linkage between economic growth of India and digital payment system RTGS is established with the help of Bayesian vector Autoregressive (BVAR) model by (Rooj & Sengupta, 2020). The monthly value and volume data of RTGS system used to analyse the possible impact on price and income level in the economy. With the usage of variables capturing financial transactions, price index, monetary policy and economic activity, they conclude that both RTGS and economic growth have a significant impact on each other. The macroeconomic variables and payment system data such as monthly ATMs, debit and credit card transactions are combined to forecast the currency demand (Raj, Bhattacharyya, Behera, John, & Talwar, 2020) and observed that weekly and monthly seasonality present in currency in circulation, and with an increase of card transactions demand for currency fell down.

Rooj and Sengupta (2021), presented time series forecasting of quarterly private consumption of people of India by using high-frequency monthly data of payment system indicators and macroeconomic variables such as personal credit and Index of Industrial Production (IIP) with mixed data sampling (MIDAS) regressions to forecast private consumption. The decisions made on payments or money in the market have macroeconomic implications and can be used by countries to make policies (Dias & Dias (2017)).

2.II. Studies in Nowcasting of Macro Economic Variables

A growing body of research uses data on electronic payment systems to nowcast several macroeconomic indicators(León & Ortega (2018), Buono, Mazzi, Kapetanios, Marcellino, & Papailias (2017), Aprigliano, Ardizzi, & Monteforte (2017), Galbraith & Tkacz(2013)). One of the studies highlighted the need to predict short-term dynamics of macroeconomic indicators in unprecedented times of COVID-19 (Chapman & Desai, 2021) and they used retail payment systems data with linear

and non-linear machine learning models to nowcast macroeconomic variables. Their model had resulted in 15% to 45% reduction in RMSE values and predicted figures were very close to official estimates during COVID-19 shock. In another study (Galbraith & Tkacz, 2018) debit card transactions and cheques clearance data (both in value and volume) were used to nowcast the GDP and retail sales growth rate. These payment systems are significant to minimize the nowcasted errors.
In the Indian context, large payments through electronic payment systems positively impact economic growth. However, the application of MIDAS regression using Indian macroeconomic data is limited. Whereas, Bentsen & Gorea (2021) show that MIDAS performs better in capturing the causal relationships between macroeconomic variables (like GDP, private and public consumption) , daily payment system data and high - frequency economic data (such as IIP, rate of unemployment, construction data, etc) than the conventional common low-frequency approach.

However, we believe that till now none of the research papers have used methods like Prophet, Prophet-with-XGBoost-errors, Feedforward neural network or their weighted ensembles and compared their efficacy with respect to traditional models like ARIMA, TBATS, ETS and their seasonal variants in the context of payment systems.

This paper is the first to use methods like Prophet, Prophet-with-XGBoost-errors, Feedforward neural network or their weighted ensembles and compared their efficacy with respect to traditional models like ARIMA, TBATS, ETS and their seasonal variants. These models examine the ability of digital payment data to nowcast GDP and consumption expenditure, focusing on India, which uses a broad set of digital payment indicators, as discussed in the later part of the paper.

2.III. Studies in use of machine learning, deep learning in Central Banks

The usage of big data is important for central banks (Sivakumar, 2018). He mentioned that big data's flexibility and real-time availability allow central banks to extract more immediate economic signals, apply new statistical approaches, improve economic predictions and financial stability assessments, and get quick feedback on policy consequences. The big data analytics and artificial intelligence can be used to track economic and financial activity by central banks (Wibisono, Ari, Widjanarti, Zulen, & Tissot, 2019). The central banks are exposed to financial big datasets which lead to large electronic footprints and financial records can be used to strengthen the economic analysis by usage of artificial intelligence and big data analytics which can outperform the traditional indicators. The big data analytics proved useful in forecasting and monitoring developments in financial markets which is a key area to central banks.

The central banks forecasting models breakdown in the wake of COVID-19 crises (Popowicz, 2021). The issue is central banks are not organized to respond to real time events and only update their forecasts every quarter. Though nowcasting techniques are used with new data sources it was not integrated into policy-making process. The central banks mostly use time-series models, semi-structural and dynamic factor models for forecasting. However, the economists of Bank of Spain, St Louis Fed and ECB continue testing new tools to model.

3. Data and Methodology

3.I Data Source

The Reserve Bank of India (RBI) publishes daily data of different payment systems viz., Real Time Gross Settlement (RTGS) system, National Electronic Funds Transfer (NEFT) system, Unified Payments Interface (UPI), National Automated Clearing House (NACH – Credit(Cr)/Debit (Dr)), Card Payments (Credit, Debit and PPI (Prepaid Payment Instruments) cards) data and cash withdrawal at Automated teller machine (ATM) and Point of sale (PoS) terminals, Cheque Truncation System (CTS) data every day on its website (www.rbi.org.in). Daily data is the best way to identify the trends and seasonality in digital payment systems. This research paper exclusively takes the daily data published by RBI. The data period of study in this paper is January 2020 to February 2022

3.II Methodology

The payment systems viz., UPI, NEFT, RTGS, NACH-Dr, Credit cards, Debit cards and IMPS are considered for forecasting. The data was partitioned into training and testing sets and some part of the training set was reserved for cross-validation. We used time series rolling cross validation for our models. Once we tuned the hyperparameters of the models, it was available for forecasting. Upon forecasting, the forecasted data were used for nowcasting the macroeconomics variable. We show that RTGS volume gives high nowcasted accuracy for GDP (at current prices) and PFCE variables. According to the Fisher's quantity theory of money

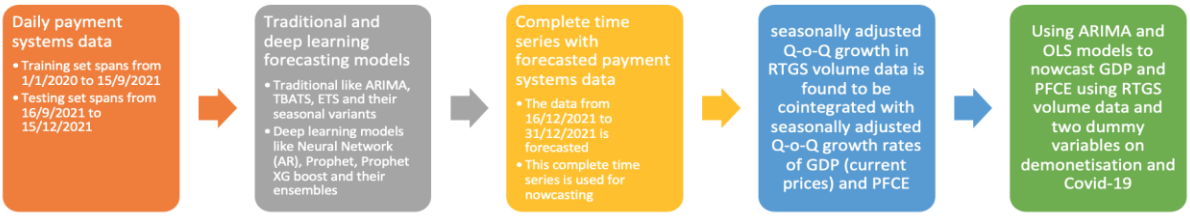
$$M \times V = P \times T$$

Where,

M=money supply; V=velocity of money; P=price level associated with transactions for the economy during the period; and T= an index of the real value of aggregate transactions.

Thus, for a given level of money supply and average price level, the real value of aggregate transactions (proxy for GDP) is directly proportional to the velocity of money. The velocity of money is more aligned to the volume of transactions compared with the value of transactions done through RTGS. Moreover, RTGS volume was difference stationary and showed cointegration with the GDP (as shown in Table 10), however, the RTGS value doesn't show these characteristics.

Figure 1: Block Diagram Showing Stages in Forecasting and Nowcasting in the Current Study



The forecasted performance was compared using two measures of errors – Root Mean Square Error (RMSE) and taking percentage difference of forecasted volumes from the actual observed volumes. We use Vector Error Correction Model (VECM) for establishing causality and long term equilibrium relation between the RTGS volume and the GDP and PFCE data. Once it is established, we use ARIMA and Ordinary Least Squares (OLS) method to nowcast the macroeconomic variables (viz., GDP and PFCE). We used the seasonally adjusted growth rates of GDP, PFCE and RTGS volume as it was first established that the seasonally adjusted time series of the same were stationary in nature (by using augmented Dickey-Fuller test), hence fulfilling the condition of cointegration.

In this paper we used both traditional and deep learning techniques to forecast payment system indicators data. The details of these techniques are explained in the following sub sections.

4. Methods and Models used in the study

4.I Traditional time series forecasting methods

Firstly, ARIMA model was estimated. ARIMA is a class of models that deal with time series data based on its previous values. These past values are the series' lags and lagged prediction errors that help to forecast the future data points using same equation. Secondly, seasonal ARIMA estimates were calculated for the payment systems under study. The model diagnostics like Autocorrelation and Partial Autocorrelation functions were observed for different permutations and combinations of models for choosing the best ARIMA model. Finally, the most accurate set of models was chosen based on AIC and BIC criteria.

Seasonal ARIMA was estimated based on the following equation -

$$\boldsymbol{\emptyset_P(B)\ \Phi_P(B^S)(1-B)^d(1-B^S)^D y_t = \theta_q(B)\Theta_Q(B^S)\varepsilon_t}$$

with

$$\emptyset_P(B) = 1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_P B^P,$$
$$\Phi_P(B) = 1 - \Phi_S B^S - \Phi_{2S} B^{2S} - \cdots - \Phi_{PS} B^{PS},$$
$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_P B^P,$$
$$\Theta_Q(B) = 1 - \Theta_S B^S - \Theta_{2S} B^{2S} - \cdots - \Theta_{QS} B^{QS},$$

where,

s = seasonal length

B = back shift operator: $B^k y_t = y_{t-k}$

ε_t = white noise sequence

$(1-B)^d$ = non-seasonal differencing operator

$(1 - B^S)^D$= seasonal differencing operator

Next, we fitted exponential smoothing model on the payment systems data mentioned above. The **Exponential Smoothing** technique uses an exponential window function to smooth data from a time series. The exponential functions, unlike the ordinary moving average, assign exponentially decreasing weights over time. The more recent values or observations are given higher weights, whereas the older values or observations are given lower weights. The exponential smoothing function is used to smooth data and operates as a low pass filter to remove high-frequency noise. A state-space ETS (Error, Trend, Seasonal) model was used. For this notation ETS (-, -, -), the possibilities for Error is {A, M}, for Trend is {N, A, Ad} and for Seasonal is {N, A, M}.

where,

A = Additive
M = Multiplicative
N = None
Ad = Damped Additive

The appropriate permutations and combinations of ETS models was chosen based on AIC criteria.

Trigonometric, Box-Cox transform, ARMA errors, Trend, and seasonal components of **TBATS** is a forecasting model based on the concept of Exponential Smoothing. TBATS model deals with complex seasonalities by modelling each one of them with Fourier series-based trigonometric representation. Such type of complex seasonality can be found in data with daily observations such as payment systems because they have both weekly and yearly seasonality.

Mathematically, the TBATS model is defined as

$$\boldsymbol{y_t^{(\lambda)} = l_{t-1} + \emptyset b_{t-1} + \sum_{i=1}^{T} s_{t-m_i}^{(i)} + d_t}$$

$$\boldsymbol{l_t = l_{t-1} + \emptyset b_{t-1} + \alpha d_t}$$

$$\boldsymbol{b_t = \emptyset b_{t-1} + \beta d_t}$$

$$\boldsymbol{d_t = \sum_{i=1}^{p} \varphi_i d_{t-i} + \sum_{i=1}^{q} \theta_i e_{t-i} + e_t}$$

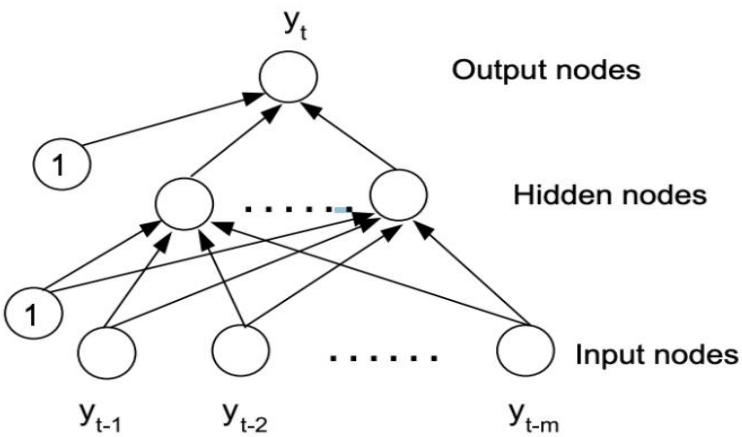
where,

$y_t^{(\lambda)}$ = *time series at t (Box-Cox transformed)*
$s_t^{(i)}$ = *ith seasonal component*
l_t = *local component*
b_t = *damping trend*
d_t = *ARMA(p, q) process for residuals*
e_t = *Gaussian white noise*

TBATS model was estimated by carefully tuning the hyperparameters of the model and using AIC method for model selection.

4.II Machine learning and Deep learning models

Feedforward Neural Network model was created for forecasting of payment system indicators. Here, Neural Network Auto-Regressive (NNAR) model was modelled. Neural networks are a type of flexible nonlinear model that learns to recognize patterns in data over time. In theory, with the right number of nonlinear processing units, such networks can learn from experience and accurately estimate any complex functional connection. For time series forecasting problems, feed-forward neural networks have proven to give high accuracy results in fairly complex scenarios. The below figure shows the feed-forward neural network with three layers. The inputs to the network are the lagged data instances of the time series, nodes in the hidden layer uses non-linear transfer functions for processing the information from the input stage, and the output layer provides the forecasted future value.



The resultant model is written as:

$$y_t = \alpha_0 + \sum_{j=1}^{n} \alpha_j f\left(\sum_{i=1}^{m} \beta_{ij}\ yt - i + \beta_{0j}\right) + \varepsilon_t$$

where,

m = no. of input nodes
n = no. of hidden nodes
f = transfer function (sigmoid) defined as

$$f(x) = \frac{1}{1+e^{-x}}$$

$\{\alpha_{j,}$ j = 0, 1, …n} is weights from hidden to output nodes.
$\{\beta_{ij,}$ i = 0, 1, …m} is weights from input to hidden nodes.
$\alpha_0\ and\ \beta_{0j}$ = arcs weights from bias terms = 1.

Apart from the above method, another newly emerging model developed by Facebook Inc. named Prophet was also used. The prophet is an open source forecasting package available in Python as well as in R. This model is optimal for various tasks having characteristics such as historical data with hourly, daily, or weekly observations, strong seasonality patterns, key holidays occurring at irregular time periods, huge outliers and missing data, various trend changes in the data or those with non-linear curves. The prophet model has an Additive Regression model procedure with four key features:

1. The trend of a piecewise linear or logistic growth curve. By picking changepoints from the data, Prophet finds changes in trends automatically.
2. Fourier Series modelled yearly seasonal component.
3. Dummy variables modelled weekly seasonal component.
4. List of key holidays provided by the user.

For forecasting, decomposable model with trend, seasonality and holidays is used, denoted by:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t$$

Where,

g(t) = trend function
s(t) = periodic changes (yearly/weekly seasonality)
h(t) = holidays effect
ε_t = error term

We modified the Prophet algorithm by using an XGBoost algorithm to learn from the mistakes of previous trees. XGBoost Algorithm is an extreme gradient boosting (ensembling) technique widely used for time series forecasting problems. Sometimes, using a single machine learning model does not give desirable results, therefore, combining various models into one and aggregating their predicting power yields a superior model with highly accurate results. However, ensemble models can have the same or different learning algorithms. The two ensemble learners Bagging and Boosting are widely used with these different statistical models with usage of decision trees. The trees are built successively in boosting, with each succeeding tree aiming to reduce the previous tree's mistakes. Each tree builds on the knowledge of its ancestors and corrects any lingering faults. As a result, the following tree in the sequence will learn from an updated set of residuals.

It may be noted that extensive hyperparameter tuning was done for the deep learning methods. Most of the deep learning methods are computationally expensive and take even hours to tune to attain desired accuracy.

4.III Ensembling technique

Different permutations and combinations of week learners were combined to become a strong learner. We ensembled the models based on their weighted averages. Weights were static and inverse of RMSE of their forecast based on testing data. Deep learning methods were combined to gain maximum possible accuracy using ensembling technique.

4.IV Nowcasting

We show that a derivative of RTGS volume is cointegrated with GDP (current prices) and PFCE. We also make use of two dummy variables. One, depicting demonetisation and second, the COVID-19 induced pandemic. First of all we establish a long term equilibrium relationship between RTGS and GDP/ PFCE using Engle and Granger as well as Johansen cointegration tests. Once established, we nowcast the GDP/PFCE using OLS and ARIMA method.

4.V Cointegration

In most cases, when the variables $\mathcal{Y}_{1,t}$ and $\mathcal{Y}_{2,t}$ are non-stationary $I(1)$ variables, a linear combination of these variables will also be non-stationary. However, in a few cases the linear combination of these variables may be stationary. This happens when the variables share the same stochastic trends, which are cancelled out when combined. In these cases, we say that the variables are cointegrated.
To see how this translates into practice, consider two variables, $\mathcal{Y}_{1,t}$ and $\mathcal{Y}_{2,t}$, which are integrated of the first order, $I(1)$. When regressing these variables on one another, we could rearrange the linear regression model, such that

$$\mathcal{U}_t = \mathcal{Y}_{1,t} - \beta_1 \mathcal{Y}_{2,t}$$

Now if the error term, $\mathcal{U}_t$ is stationary, $I(0)$, then by definition the combined $\mathcal{Y}_{1,t} - \beta_1 \mathcal{Y}_{2,t}$ must also be stationary since the properties of the left-hand-side must equal the properties on the right-hand-side. Hence, while both $\mathcal{Y}_{1,t}$ and $\mathcal{Y}_{2,t}$ have stochastic trends, we say that the variables $\mathcal{Y}_{1,t}$ and $\mathcal{Y}_{2,t}$ are cointegrated, as the linear combination $\mathcal{Y}_{1,t} + \beta_1 \mathcal{Y}_{2,t}$ has the same statistical properties as an $I(0)$ variable. Note that these stochastic trends are related through β_1, which contains this feature (relating to the common stochastic trends) of the data. Of course, if $\mathcal{U}_t$ is non-stationary, $I(1)$, as would usually be the case, then $\mathcal{Y}_{1,t}$ and $\mathcal{Y}_{2,t}$ are not cointegrated and regressing $\mathcal{Y}_{2,t}$ on $\mathcal{Y}_{1,t}$ would yield a spurious result. Where $\boldsymbol{y}_t = \left(\mathcal{Y}_{1,t}, \mathcal{Y}_{2,t}\right)'$ is a (2 x 1) vector of $I(1)$ variables. Then the coefficient matrix may be given as, $\beta = (1 - \beta_1)'$, where the relationship between the variables could be summarized as $\beta' \boldsymbol{y}_t = \mathcal{Y}_{1,t} - \beta_1 \mathcal{Y}_{2,t}$. These variables will then be cointegrated when $\beta' \boldsymbol{y}_t \sim I(0)$. The linear combination of variables to derive $\beta' \boldsymbol{y}_t$ will typically be motivated by economic theory and this often referred to as the long-run equilibrium relationship. The idea is that variables in the $\boldsymbol{y}_t$ vector will drift together as they follow some form of long-run equilibrium. The vector β is termed the cointegrating vector, which summarizes the relationship between the stochastic trends. When components of $\boldsymbol{y}_t$ are integrated of order d and the reduction in the order of the combined variables is b. then we note that $\boldsymbol{y}_t \sim CI\ (d, b)$.

4.VI Error Correction Models

Where a cointegrating relationship may be used to define an equilibrium relationship, the time paths of cointegrated variables are influenced by the extent of any deviation from the long run equilibrium. If the variables are cointegrated then they will return towards the equilibrium values, although they need not actually attain these values at a particular point in time. What is essential is that there is a force that will draw the variables towards the equilibrium values, so that the deviation from equilibrium (following a shock) is not permanent.
The deviation of a cointegrated variable from the path of equilibrium may be modelled with the aid of an error correction representation. Engle and Granger formalized the connection between this dynamic response to the errors and co-integration in the Engle-Granger representation theorem, which states that two variables are cointegrated if, and only if, there exists an error correction mechanism for one set of variables.

4.VII Vector Error Correction Representation

The general form of the error correction mechanism allows for richer dynamics interactions between the variables, which can be specified as,

$$\Delta \mathcal{Y}_{1,t} = \gamma_0 + \alpha_1 \left[\mathcal{Y}_{1,t-1} - \beta_1 \mathcal{Y}_{2,t-1}\right] + \sum_{i=1}^{K} \zeta_{1,i}\, \Delta \mathcal{Y}_{1,t-1} + \sum_{j=1}^{L} \zeta_{2,i}\, \Delta \mathcal{Y}_{2,t-1} + \varepsilon_{\mathcal{Y}_{1,t}}$$

$$\Delta \mathcal{Y}_{2,t} = \eta_0 + \alpha_2 \left[\mathcal{Y}_{1,t-1} - \beta_1 \mathcal{Y}_{2,t-1}\right] + \sum_{i=1}^{K} \zeta_{1,i}\, \Delta \mathcal{Y}_{2,t-1} + \sum_{j=1}^{L} \zeta_{2,i}\, \Delta \mathcal{Y}_{1,t-1} + \varepsilon_{\mathcal{Y}_{2,t}}$$

This representation is termed the vector error correction model (VECM), where it would be possible to show that if both α_1 and α_2 are equal to zero, then there is: no equilibrium relationship, no error-correction, and no cointegration. Note that if $\mathcal{Y}_{1,t}$ and $\mathcal{Y}_{2,t}$ are $CI(1,1)$ then all terms in either of the above equations are $I(0)$ and statistical inference that makes use of standard t and F statistics would be applicable.

5. Empirical results

Forecasting results for select payment systems are illustrated below.

Table 3: Relative Forecasting Accuracy of Different Models

Model	RTGS		UPI		NEFT		IMPS		Credit Card		Debit Card	
	RMSE	Error*	RMSE	Error*	RMSE	Error*	RMSE	Error*	RMSE	Error*	RMSE	Error*
ARIMA XGBOOST	2.69	2.25	69.16	-2.33	44.39	-1.24	12.67	-4.96	8.01	-9.15	8.96	-0.81
TBATS	1.84	-7.22	79.23	-2.83	36.65	-5.09	11.47	-4.09	5.44	-12.39	5.46	0.79
SEASONAL ETS	1.82	-7.72	84.18	-3.58	39.71	-13.28	12.77	-5.49	5.56	-11.58	5.96	0.90
SIMPLE ETS	1.84	-6.67	81.44	-2.76	34.15	3.59	11.67	-4.45	5.13	-11.45	5.53	1.18
NNETAR	1.79	-3.43	293.54	-18.51	30.30	0.41	10.30	-0.30	4.93	-3.27	7.14	12.74
SIMPLE PROPHET	2.04	7.13	90.06	-2.47	35.40	-1.73	9.72	4.47	5.28	-0.32	7.93	12.53
ProphetXGBOOST	1.86	-0.67	67.37	3.08	28.75	-16.37	5.62	1.74	5.12	-1.50	6.42	0.05
WEIGHTED MEAN	1.81	0.77	61.45	0.70	26.69	-6.30	5.93	1.96	4.92	-1.73	5.29	0.75

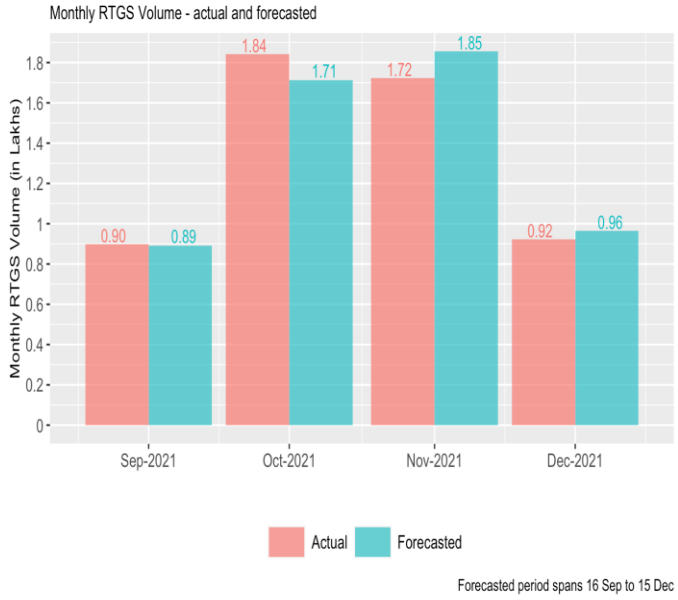
Note: Error is calculated as 100(forecasted value – actual value)/ actual value

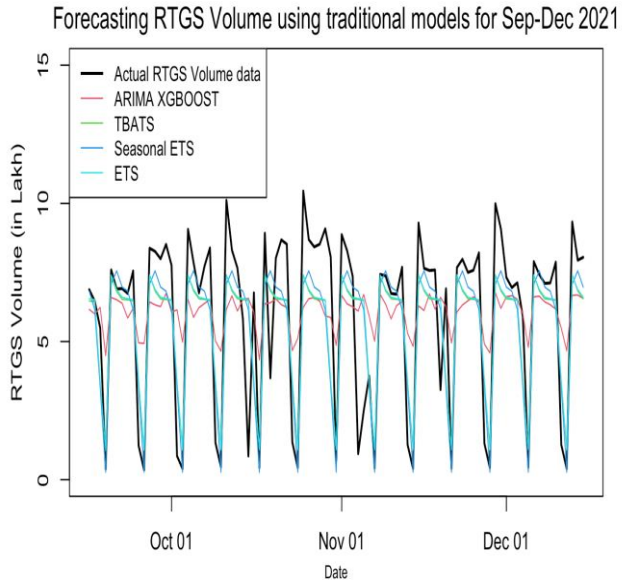
5.I RTGS volume

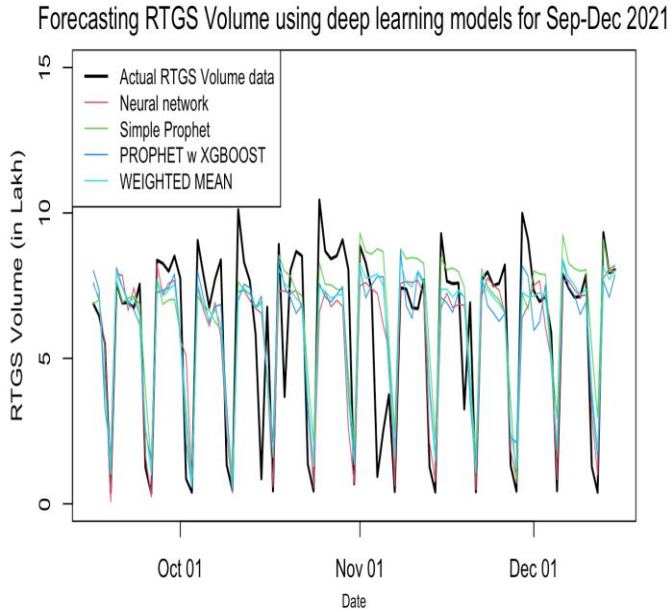
Table 4: Performance of Different Forecasting Models for RTGS Volume

RTGS Volume - Error metrics for Sep-Dec 2021 forecasting

	RMSE	ERROR
ARIMA XGBOOST	2.692	2.253
TBATS	1.844	-7.215
SEASONAL ETS	1.820	-7.719
SIMPLE ETS	1.836	-6.674
NNETAR	1.793	-3.427
SIMPLE PROPHET	2.037	7.134
PROPHET XGBOOST	1.856	-0.673
WEIGHTED MEAN	1.812	0.774







5.II UPI volume

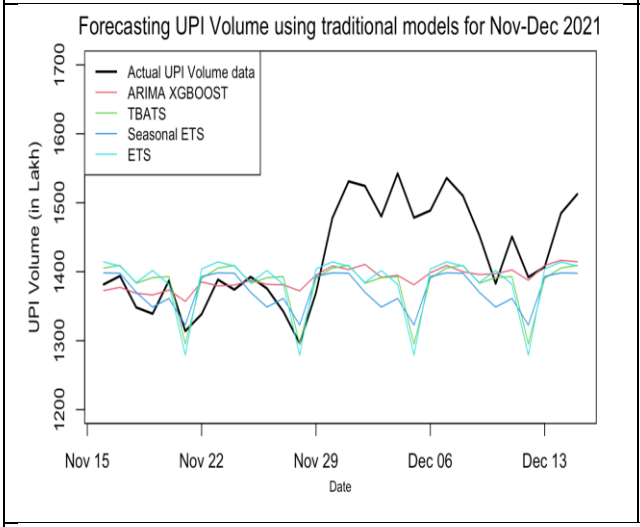
Table 5: Performance of Different Forecasting Models for UPI Volume

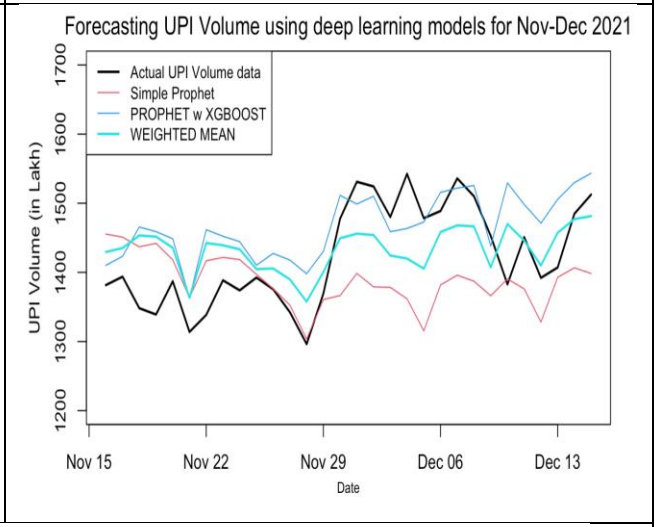
UPI Volume - Error metrics for Nov-Dec 2021 forecasting

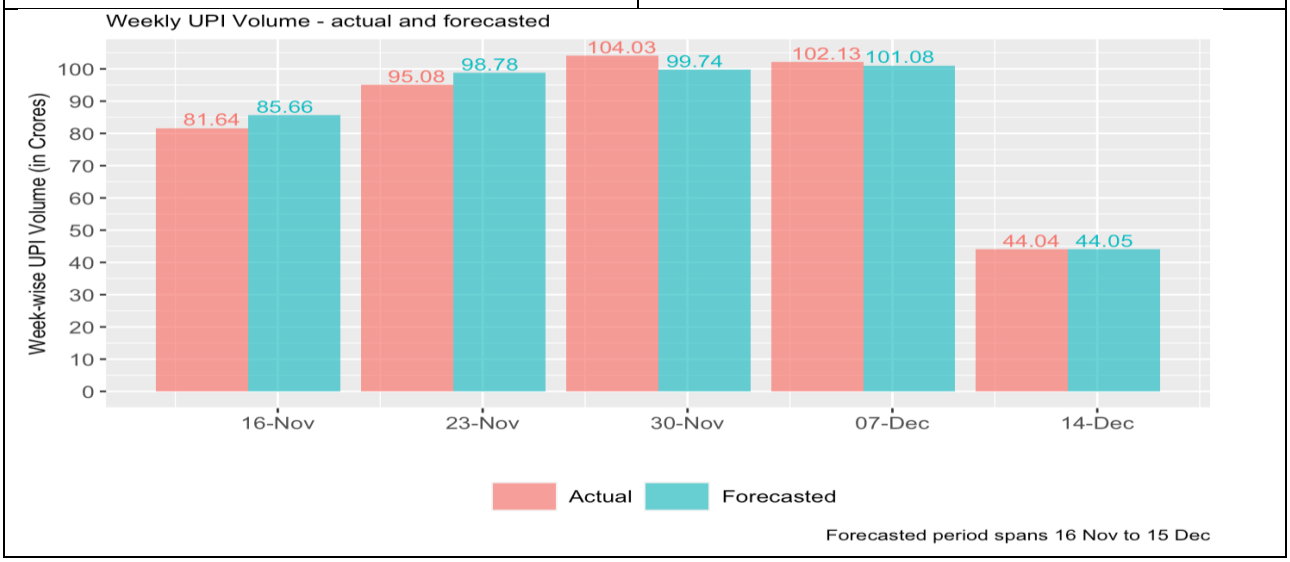
	RMSE	ERROR
ARIMA XGBOOST	69.164	-2.328
TBATS	79.231	-2.828
SEASONAL ETS	84.181	-3.579
SIMPLE ETS	81.436	-2.757
NNETAR	293.543	-18.514
SIMPLE PROPHET	90.061	-2.470
PROPHET XGBOOST	67.371	3.078
WEIGHTED MEAN	61.454	0.703

UPI Volume - Week-wise forecasting accuracy

Week	Weighted Mean	Real_UPI_Vol	Error %
2021-11-15	8568.53	8163.70	5.0
2021-11-22	9872.88	9507.92	3.8
2021-11-29	10009.79	10403.46	-3.8
2021-12-06	10126.11	10213.32	-0.9
2021-12-13	4415.90	4404.47	0.3





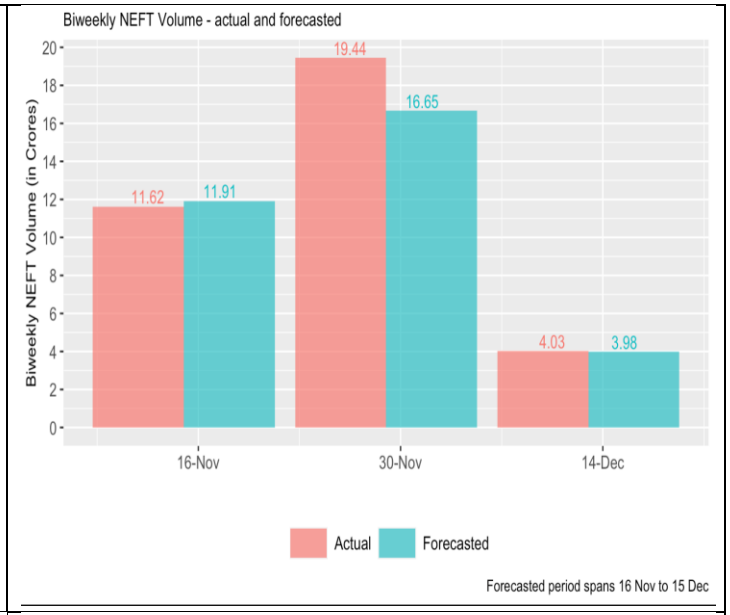


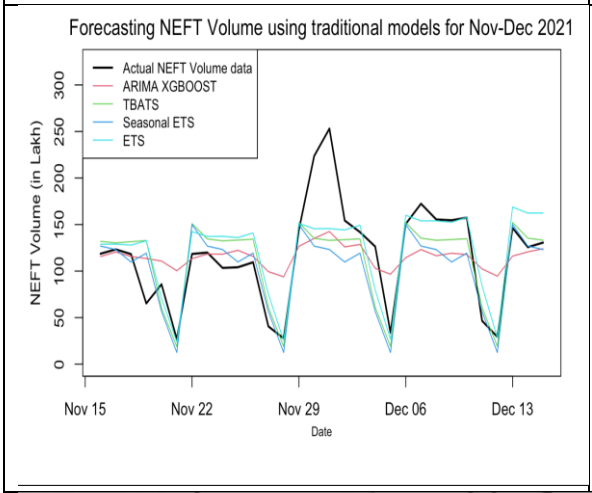
5.III NEFT volume

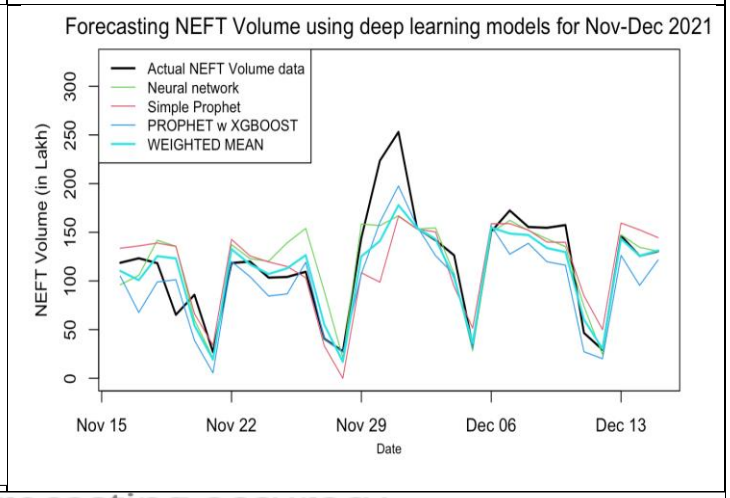
Table 6: Performance of Different Forecasting Models for NEFT Volume

NEFT Volume - Error metrics for Nov-Dec 2021 forecasting

	RMSE	ERROR
ARIMA XGBOOST	44.392	-1.236
TBATS	36.645	-5.094
SEASONAL ETS	39.705	-13.284
SIMPLE ETS	34.152	3.588
NNETAR	30.298	0.406
SIMPLE PROPHET	35.395	-1.727
PROPHET XGBOOST	28.753	-16.372
WEIGHTED MEAN	26.694	-6.298







NEFT Volume - Biweekly forecasting accuracy

Week	Weighted Mean	Real_NEFT_Vol	Error %
2021-11-15	1190.94	1161.60	2.53
2021-11-29	1665.38	1943.71	-14.32
2021-12-13	397.72	402.64	-1.22

5.IV IMPS volume

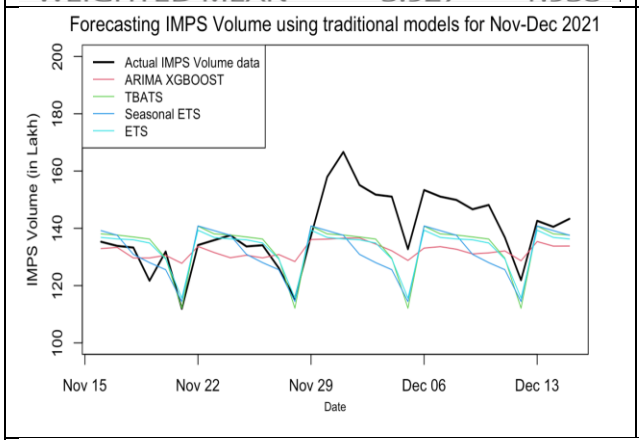
Table 7: Performance of Different Forecasting Models for IMPS Volume

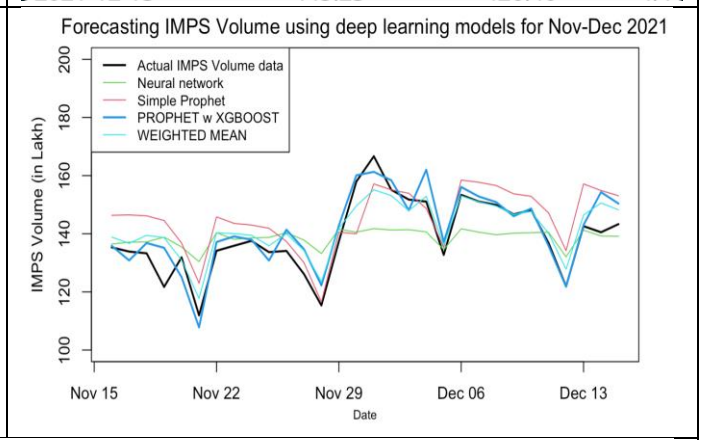
IMPS Volume - Error metrics for Nov-Dec 2021 forecasting

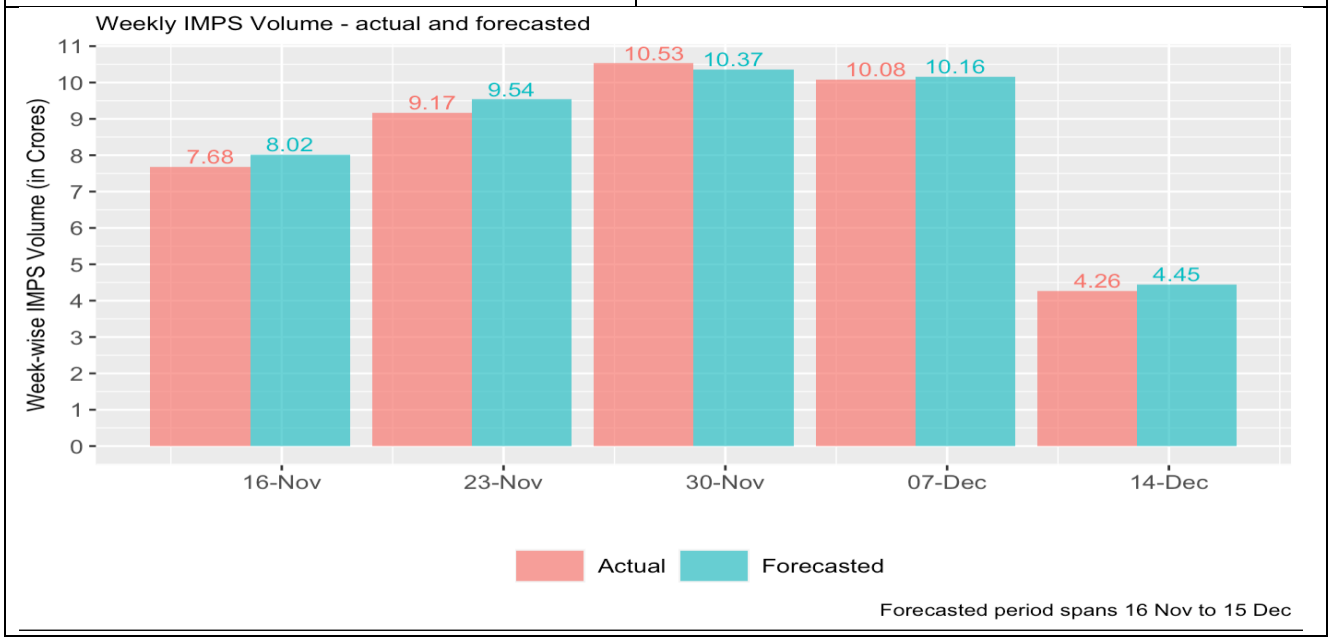
	RMSE	ERROR
ARIMA XGBOOST	12.672	-4.955
TBATS	11.471	-4.089
SEASONAL ETS	12.767	-5.486
SIMPLE ETS	11.673	-4.449
NNETAR	10.298	-0.304
SIMPLE PROPHET	9.716	4.465
PROPHET XGBOOST	5.623	1.742
WEIGHTED MEAN	5.927	1.958

IMPS Volume - Week-wise forecasting accuracy

Week	Weighted Mean	Real_IMPS_Vol	Error %
2021-11-15	802.41	767.98	4.5
2021-11-22	953.52	916.89	4.0
2021-11-29	1036.54	1052.62	-1.5
2021-12-06	1015.82	1007.89	0.8
2021-12-13	445.23	426.46	4.4

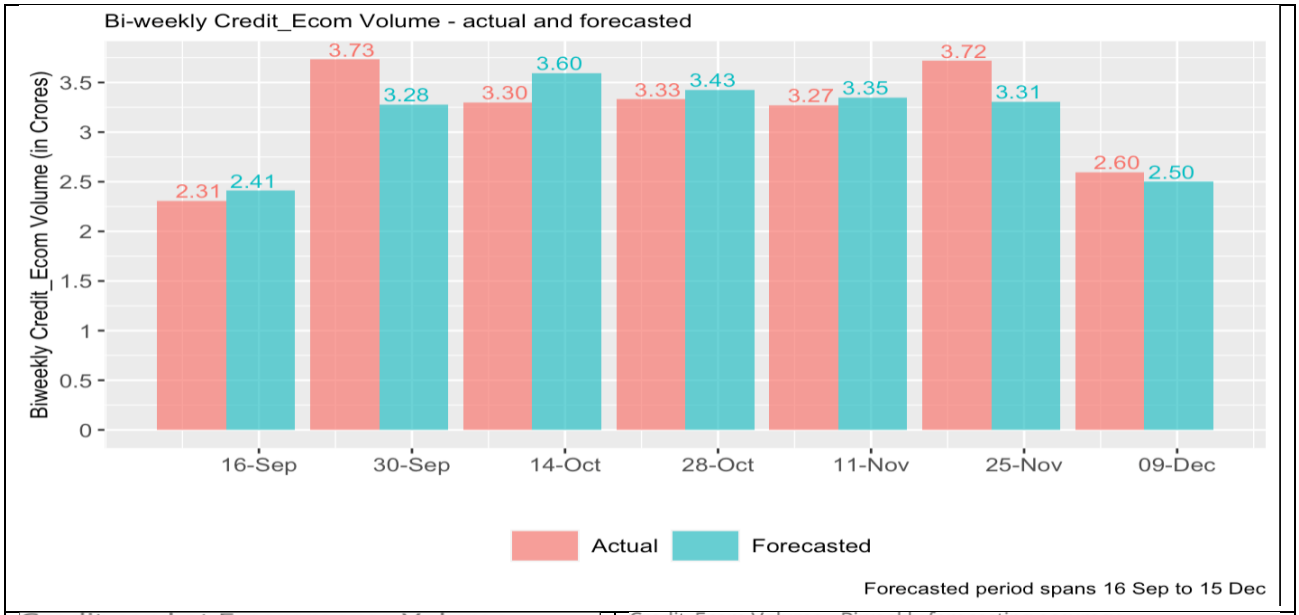






5.V Credit card volume

Table 8: Performance of Different Forecasting Models for Credit Card Volume

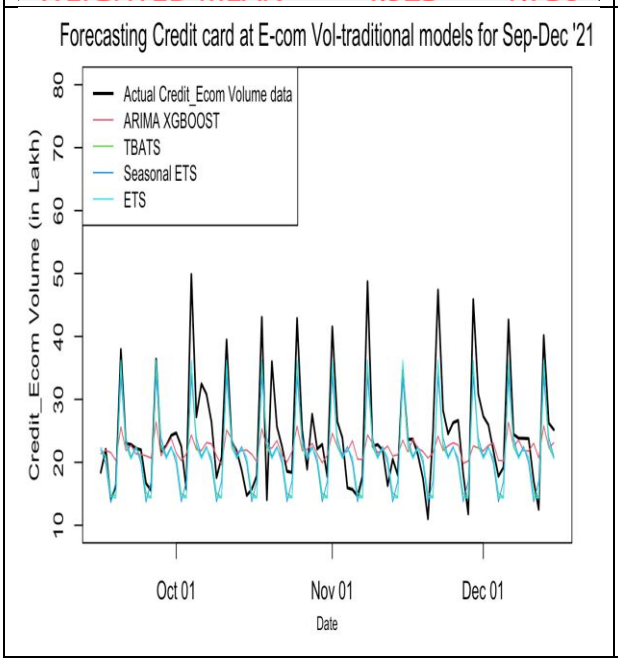


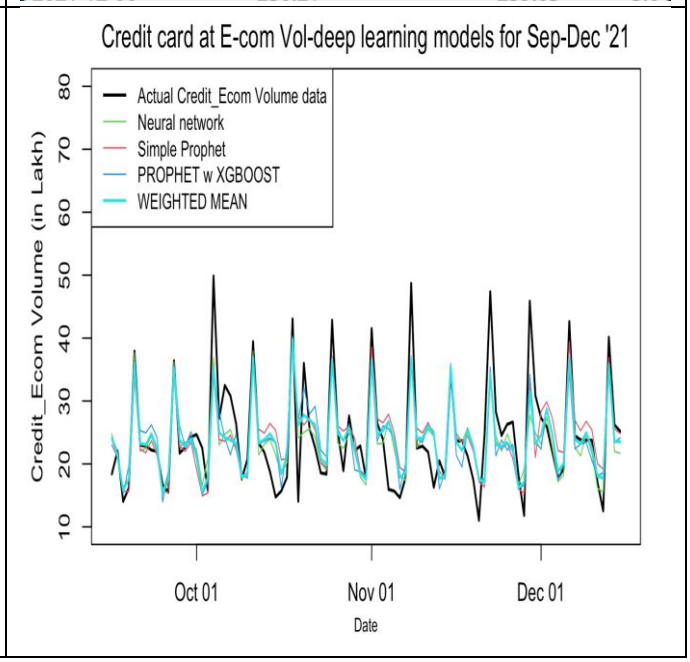
Credit card at E-commerce Volume - Error metrics for Sep-Dec 2021 forecasting

	RMSE	ERROR
ARIMA XGBOOST	8.011	-9.153
TBATS	5.444	-12.385
SEASONAL ETS	5.559	-11.579
SIMPLE ETS	5.132	-11.448
NNETAR	4.929	-3.267
SIMPLE PROPHET	5.276	-0.321
PROPHET XGBOOST	5.119	-1.501
WEIGHTED MEAN	4.923	-1.730

Credit_Ecom Volume - Biweekly forecasting accuracy

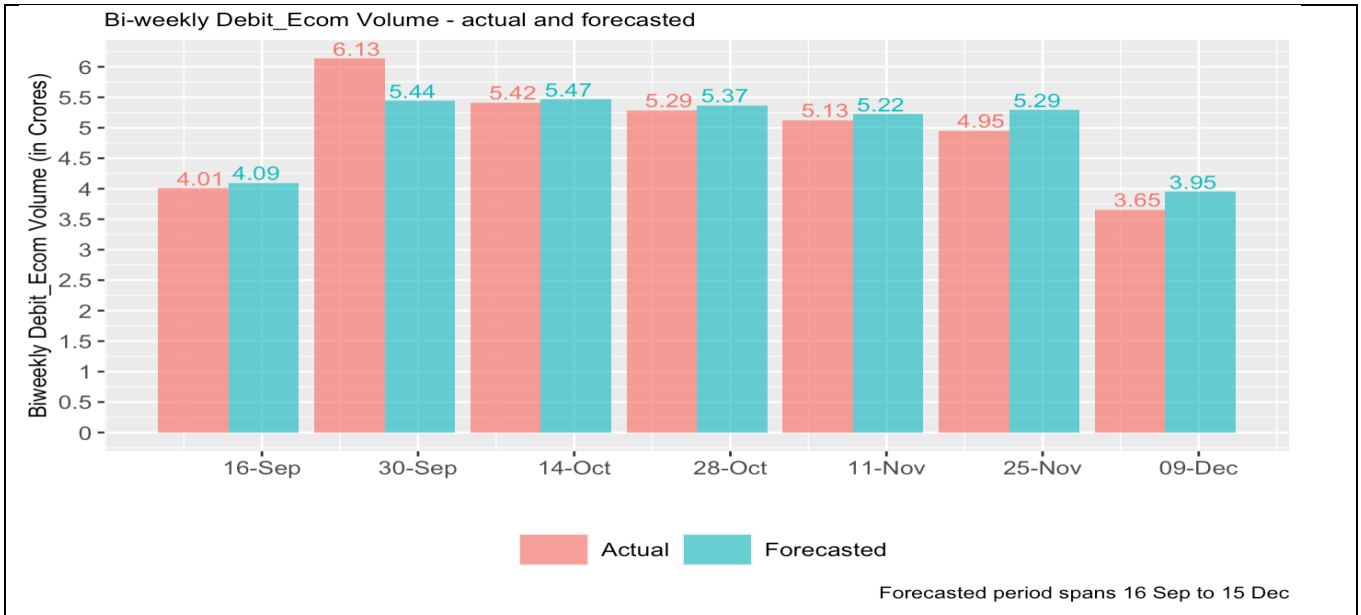
Week	Weighted Mean	Real_Credit_Ecom_Vol	Error %
2021-09-13	241.10	230.83	4.4
2021-09-27	327.97	373.11	-12.1
2021-10-11	359.59	329.87	9.0
2021-10-25	342.62	333.24	2.8
2021-11-08	335.09	327.22	2.4
2021-11-22	330.75	371.90	-11.1
2021-12-06	250.21	259.63	-3.6





5.VI Debit card volume

Table 9: Performance of Different Forecasting Models for Debit Card Volume

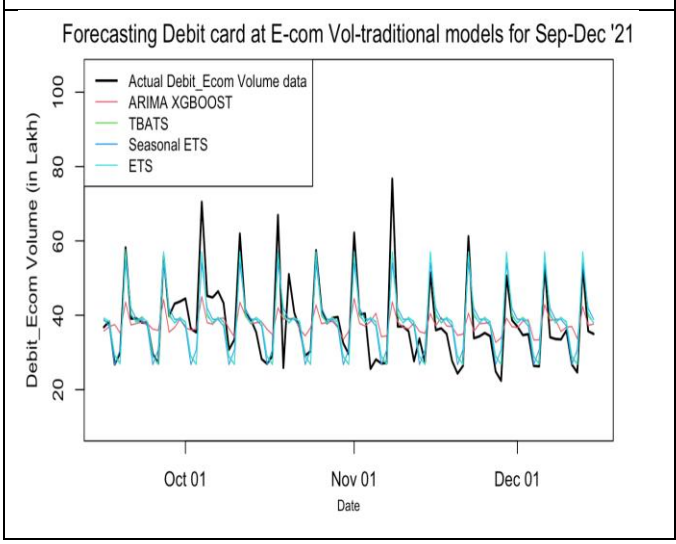


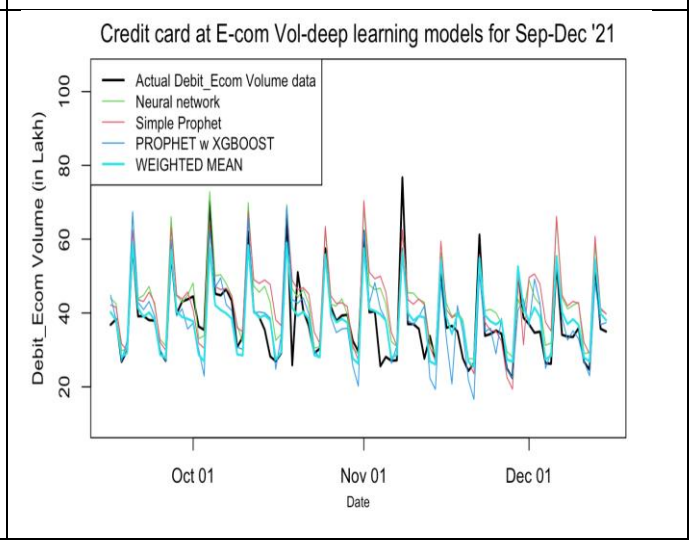
Debit card at E-commerce Volume - Error metrics for Sep-Dec 2021 forecasting

	RMSE	ERROR
ARIMA XGBOOST	8.955	-0.805
TBATS	5.456	0.788
SEASONAL ETS	5.956	0.901
SIMPLE ETS	5.531	1.184
NNETAR	7.139	12.737
SIMPLE PROPHET	7.934	12.530
PROPHET XGBOOST	6.421	0.049
WEIGHTED MEAN	5.289	0.752

Debit_Ecom Volume - Biweekly forecasting accuracy

Week	Weighted Mean	Real_Debit_Ecom_Vol	Error %
2021-09-13	409.37	400.82	2.1
2021-09-27	544.16	613.34	-11.3
2021-10-11	546.68	541.58	0.9
2021-10-25	536.67	528.81	1.5
2021-11-08	522.31	512.51	1.9
2021-11-22	529.30	494.80	7.0
2021-12-06	394.86	365.46	8.0



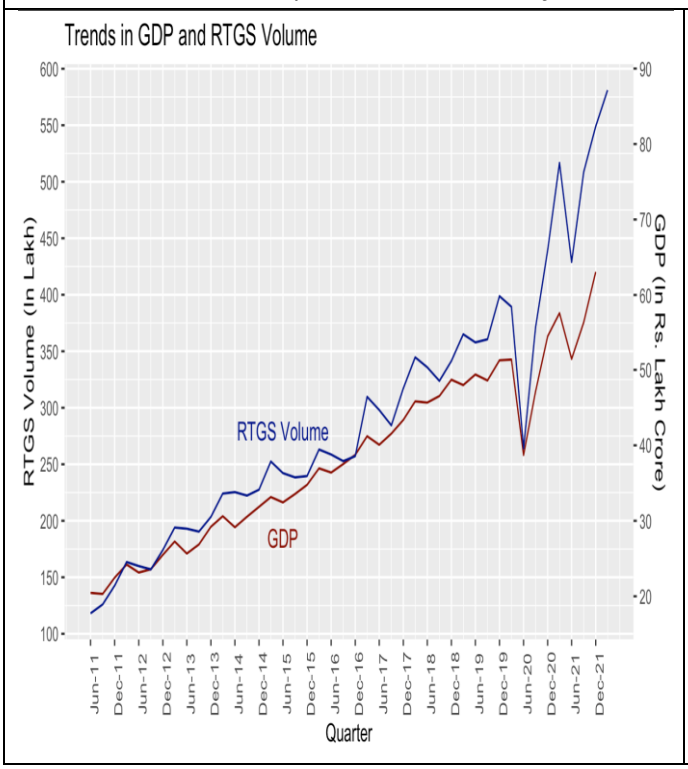


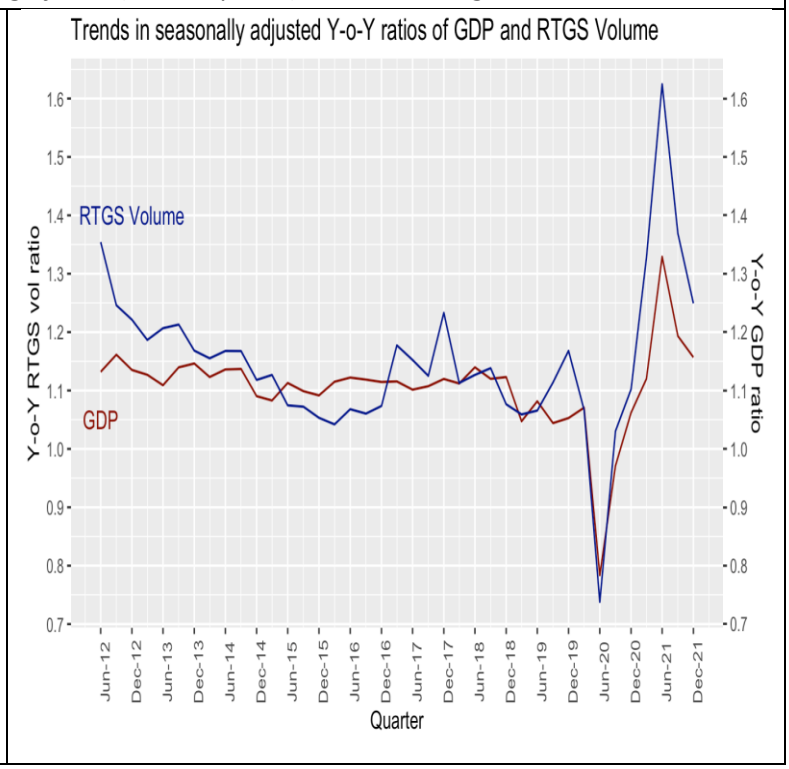
6. Nowcasting of Macroeconomic Variables

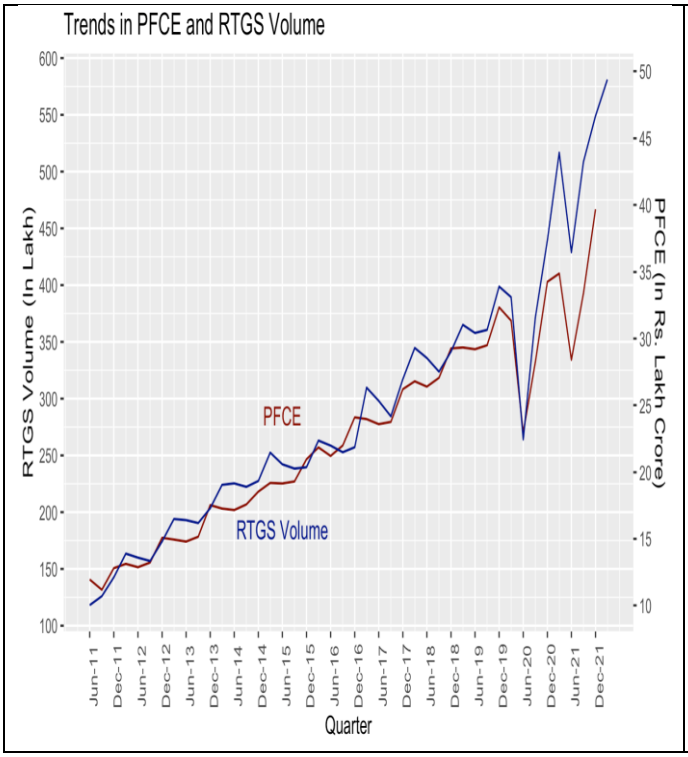
The growth of GDP (at current prices) and RTGS are made seasonally adjusted by taking ratio of the recent quarter with the quarter at an year lag. So, December 2021 quarter is divided by December 2020 quarter to calculate the seasonally adjusted Q-o-Q growth rate.

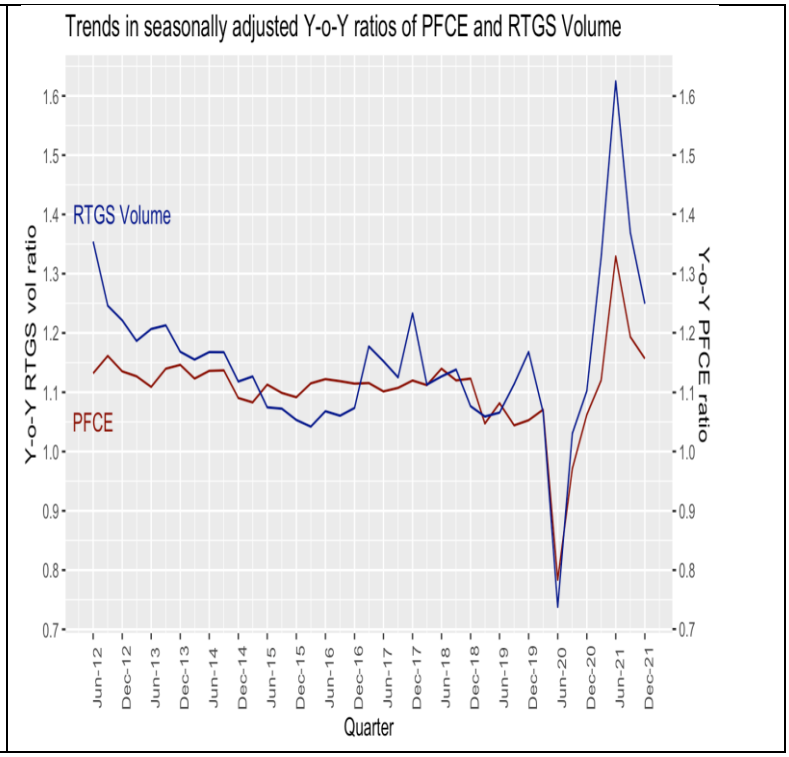
Prime facie, there appears to be a visually apparent correlation between both GDP (at current prices) and Private Final Consumption Expenditure (PFCE) with the correlation between their seasonally adjusted Q-o-Q growth rates being 0.79 and 0.69 respectively. This is further exemplified in the graphs below (Table 10).

Table 10: Graphical Presentation of Nowcasting of GDP (current prices) and PFCE using RTGS Volume









Clearly, there is some cointegration between GDP, PFCE and RTGS volume. To empirically prove it, we conducted Augmented Dickey Fuller (ADF) tests and then conducted VECM analysis on the same. The results of ADF tests are in table below –

Table 11: Augmented Dickey Fuller Tests Results for Time Series

Seasonally adjusted and differenced growth rates of variables	Dickey-Fuller statistic	p-value	Inference
GDP (at current prices)	-5.66	0.01	Stationary
PFCE	-5.32	0.01	Stationary
RTGS volume	-5.24	0.01	Stationary

The Vector Error Correction Model equation for GDP is as follows –

$$\Delta \mathrm{GDP}^4{}_t = \gamma_0 + \alpha_1 \left[\mathrm{GDP}_{t-1} - \beta_1 \mathrm{RTGS}_{t-1}\right] + \sum_{i=1}^{K} \zeta_{1,i}\, \Delta \mathrm{GDP}_{t-1} + \sum_{j=1}^{L} \zeta_{2,i}\, \Delta \mathrm{RTGS}_{t-1} + \varepsilon_{\mathrm{GDP}_t}$$

$$\Delta \mathrm{PFCE}^5{}_t = \gamma_0 + \alpha_1 \left[\mathrm{PFCE}_{t-1} - \beta_1 \mathrm{RTGS}_{t-1}\right] + \sum_{i=1}^{K} \zeta_{1,i}\, \Delta \mathrm{PFCE}_{t-1} + \sum_{j=1}^{L} \zeta_{2,i}\, \Delta \mathrm{RTGS}_{t-1} + \varepsilon_{\mathrm{PFCE}_t}$$

Where α_1 and α_2 are the Error Correction Terms (ECTs) respectively for GDP and PFCE respectively

[4] GDP is actually seasonally adjusted Q-o-Q ratio of GDP (at current prices)
[5] PFCE is actually seasonally adjusted Q-o-Q ratio of PFCE
The process of making data seasonally adjusted is explained in the above section.

Table 12: Results of VECM Regression Models for GDP and PFCE

Regression models for VECM of GDP		Regression models for VECM of PFCE	
	VECM for GDP		**VECM for PFCE**
Predictors	*Estimates*	*Predictors*	*Estimates*
Intercept	-0.01 (0.01)	Intercept	-0.01 (0.01)
1st lag of diff(RTGS)	0.59 ** (0.19)	1st lag of diff(RTGS)	0.54 ** (0.15)
2nd lag of diff(RTGS)	0.16 (0.23)	2nd lag of diff(RTGS)	0.24 (0.15)
3rd lag of diff(RTGS)	0.08 (0.26)	3rd lag of diff(RTGS)	0.30 (0.16)
4th lag of diff(RTGS)	-0.09 (0.20)	4th lag of diff(RTGS)	-0.18 (0.14)
Error correction term (ECT)	1.12 ** (0.39)	Error correction term (ECT)	1.67 *** (0.29)
1st lag of diff(GDP)	-1.11 ** (0.33)	1st lag of diff(PFCE)	-1.21 *** (0.27)
2nd lag of diff(GDP)	-0.41 (0.40)	2nd lag of diff(PFCE)	-0.68 * (0.25)
3rd lag of diff(GDP)	-0.50 (0.44)	3rd lag of diff(PFCE)	-0.88 ** (0.27)
4th lag of diff(GDP)	-0.51 (0.30)	4th lag of diff(PFCE)	-0.35 (0.19)
Observations	34	Observations	34
* $p<0.05$ ** $p<0.01$ *** $p<0.001$		* $p<0.05$ ** $p<0.01$ *** $p<0.001$	

We see that the error correction term (ECT) is significant in both GDP as well as PFCE VECM equations. Hence, there is long run relationship between the GDP/ PFCE and RTGS volume. Thus, RTGS volume can be helpful in nowcasting them.

Next we used ARIMA and OLS models to nowcast GDP and PFCE. We also used the dummy variables for the macroeconomic shocks like demonetisation and the onset of COVID-19 induced pandemic.The results are depicted in Table 13. We observe that on an average ARIMA model gives a better estimate for GDP as well as PFCE. The error per cent (measured as per cent deviation from the actual figures) were generally less than 3%. The graphs of estimated figures neatly coincided on that of the macroeconomic variables. On conducting the Forecasting Error Variance Decomposition analysis, we observe that the variation in both GDP and PFCE can be explained by RTGS volume, more for GDP than

PFCE. Also, studying the impulse response function, a sudden shock to RTGS volume would give perpetuate to the macroeconomic variables as well. We observe that the effect of shock to RTGS doesn't die down even in the long run and causes a structural shift in the GDP growth rate. Also, GDP is more sensitive to shocks to RTGS volume. A one standard deviation shock to RTGS leads to increase in GDP by upto 1% in the short run. The effect on PFCE is on the other hand just above 0.4%.

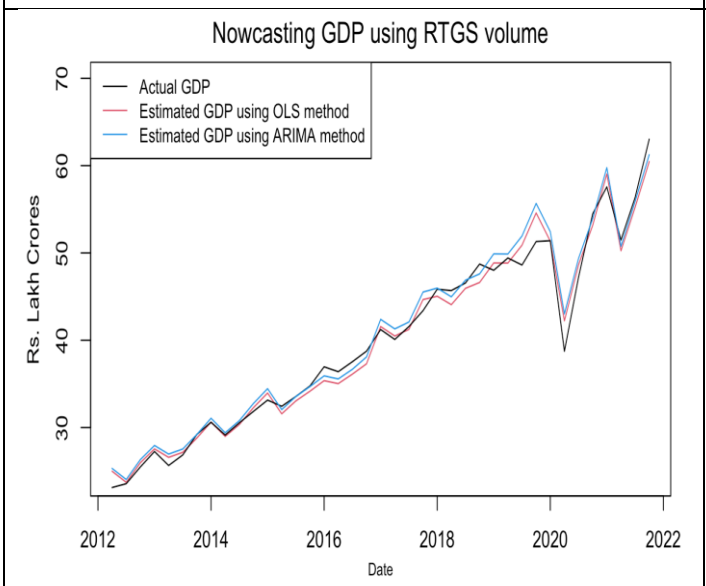
Table 13: Comparison of Performance of OLS and ARIMA Models for GDP and PFCE Nowcasting

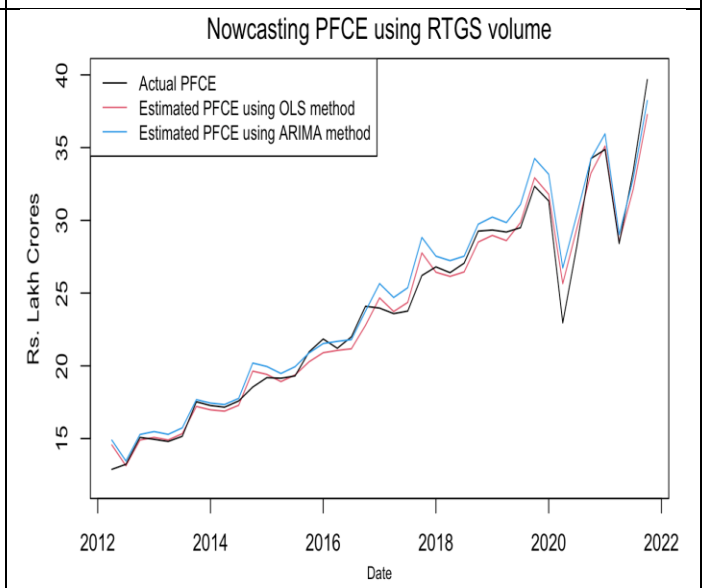
Nowcasting GDP (in Rs. Lakh Crores) using OLS and ARIMA

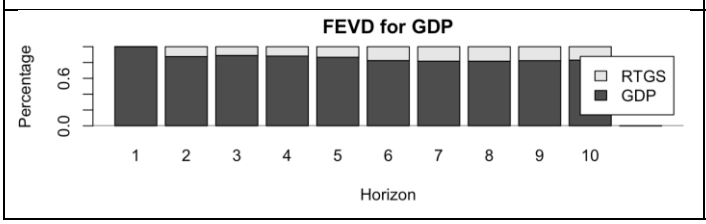
CY Qtr	Actual GDP	GDP using OLS	% error OLS	GDP using ARIMA	% error ARIMA
2019-Q3	48.62	50.89	**4.68**	51.94	**6.84**
2019-Q4	51.30	54.58	**6.39**	55.67	**8.52**
2020-Q1	51.40	51.35	**-0.11**	52.44	**2.01**
2020-Q2	38.73	42.23	**9.01**	43.02	**11.05**
2020-Q3	47.22	48.64	**3.02**	49.38	**4.57**
2020-Q4	54.49	53.16	**-2.43**	53.93	**-1.02**
2021-Q1	57.57	59.04	**2.55**	59.77	**3.82**
2021-Q2	51.48	50.25	**-2.40**	50.76	**-1.40**
2021-Q3	56.35	55.22	**-2.00**	55.89	**-0.82**
2021-Q4	63.03	60.47	**-4.07**	61.25	**-2.82**

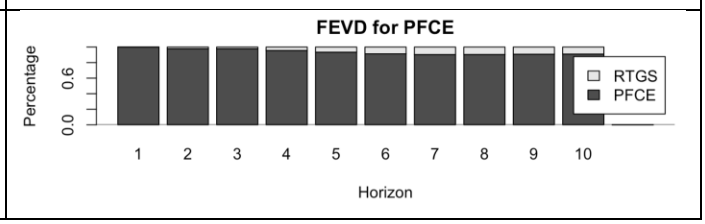
Nowcasting PFCE (in Rs. Lakh Crores) using OLS and ARIMA

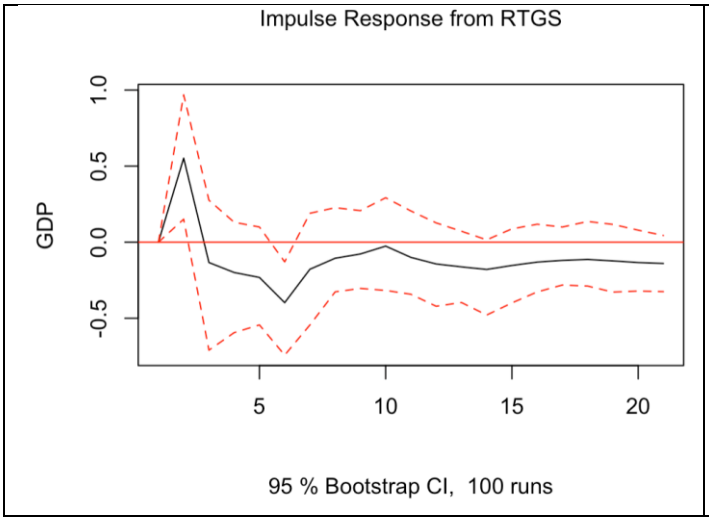
CY Qtr	Actual PFCE	PFCE using OLS	% error OLS	PFCE using ARIMA	% error ARIMA
2019-Q3	29.50	29.85	**1.19**	31.10	**5.42**
2019-Q4	32.34	32.93	**1.82**	34.25	**5.91**
2020-Q1	31.34	31.79	**1.46**	33.17	**5.85**
2020-Q2	22.96	25.65	**11.74**	26.72	**16.40**
2020-Q3	28.25	29.46	**4.28**	30.40	**7.60**
2020-Q4	34.25	33.25	**-2.92**	34.23	**-0.03**
2021-Q1	34.87	35.11	**0.67**	35.95	**3.08**
2021-Q2	28.41	28.52	**0.41**	29.02	**2.18**
2021-Q3	33.40	32.13	**-3.78**	32.87	**-1.57**
2021-Q4	39.68	37.27	**-6.06**	38.24	**-3.63**

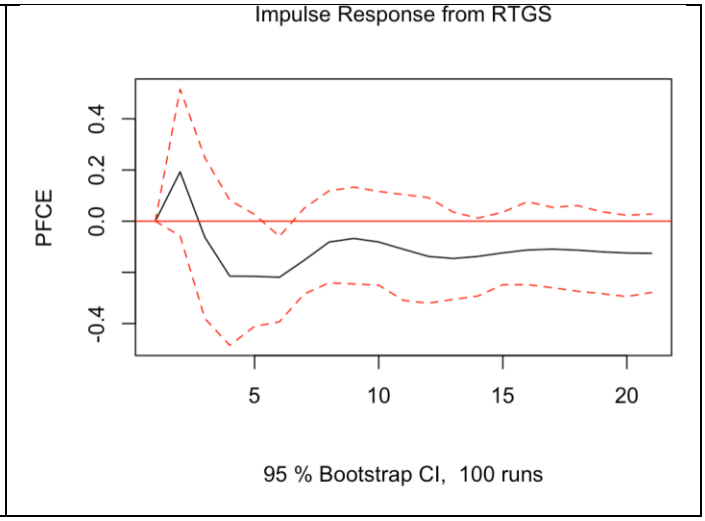












Based on the ARIMA and OLS methods, we nowcast the GDP (at Current Prices) for the March 2022 quarter is likely to be ₹ 61.77 lakh crore which resulted that the GDP in Q4 of FY 2021 is likely to increase (in Y-o-Y terms) by 7.3%. Further, based on the ARIMA and OLS models, we nowcast the Private Final Consumption Expenditure (PFCE) for the quarter ending March 2022 is likely to be ₹ 37.97 lakh crore which predicts that the PFCE in Q4 of FY 2021 is likely to increase (in Y-o-Y terms) by 8.9%.

7. Conclusions and Policy Recommendation

Forecasting of payment systems data opens up a new avenue for nowcasting of macroeconomic variables. The policy makers can keep a close tab on the unfolding picture on real economy, almost on a real time basis. Both the central bank as well as the government can take necessary actions proactively if the payment systems transactions show decline. However, even though the accuracy of nowcasting is high, the forecast can be made more consistent by including real economy variables. Bhadury, Ghosh, Kumar (2000) have used Dynamic Factor Models comprising of 28 real economy wide indicators which gives better RMSE. However, using payment system indicators provide two fold advantages. Firstly, as the data on payment system indicators are available on a daily basis, hence forecasting of remainder of quarterly data is possible with very high certainty. Thus, nowcasting of macroeconomic variables (viz., GDPand PFCE) data can be done even before the end of the quarter. Secondly, as the payment systems data is available on a daily basis, hence an economy-wide macro picture can be ascertained if one knows about the underlying dynamics between the payment systems data and the real economy macroeconomic data.

The forecasting of payment systems data has other policy implications as well. With a high accuracy of forecasted figures, a cyclical and seasonal adjustment can be made to the servers of entities which handle huge quantities of payments traffic volume. Moreover, the immediate impact of a policy change can be measured and its likely implication on the other macroeconomic variables can also be estimated using above techniques. Also, knowing the seasonality pattern of different payment system products also gives an insight into the consumer payments behaviour around significant events like festivals, holiday seasons, etc.

While the paper provided use of only payment system data in nowcasting of GDP/PCFC with high accuracy, it did not examine the results of combining payment system indicators and traditional real economy wide indicators being used for forecasting and nowcasting of GDP/PCFC due to unavailability of high frequency data of these indicators. The authors will keep on improving upon the above work by adding more relevant variables after careful analysis along with payment system indicators. Further, an avenue which may be explored is using Mixed Data Sampling approach (MIDAS) and State Space Models like Dynamic Factor Model (DFM) for nowcasting.

REFERENCES

Das, S. (2021). Financial Sector in the New Decade (*Address by Shri Shaktikanta Das, Governor, Reserve Bank of India - Thursday, March 25, 2021 - at the Times Network India Economic Conclave 2021 in New Delhi*)

Reserve Bank of India (2021). *Booklet on Payment Systems*.

Reserve Bank of India (2019). *Payment and Settlement Systems in India: Vision 2019 – 2021*

Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and Practice. (2nd ed.) OTexts. https://otexts.org/fpp2/

Acharya, V., Bhadury, S., & Surti, J. (2020). Financial Vulnerabilities and Risk to Growth in India.

Bhadury, S., Pohit, S., & Beyer, R. (2018). A New Approach to Nowcasting Indian Gross Value Added. NCAER Working Paper No.115, National Council of Applied Economic Research.

Banerjee, D. (2014). Forecasting of Indian Stocl(Market using Time-series ARIMA Model. *2014 2nd International Conference on Business and Information Management (ICBIM)*, (pp. 131-135).

Chaudhuri, T. D., & Ghosh, I. (2016). Artificial Neural Network and Time Series Modeling Based Approach to Forecasting the Exchange Rate in a Multivariate Framework. *Journal of Insurance and Financial Management, 1*(5), 92-123.

Corstjens, M., Bakhshandeh, M., Kahraman, P., & Bosman, J. (2019). Predicting the daily number ofpayment transactions in the largest bank in the Netherlands: Application to Banking Data. *2019 IEEE International Conference on Big Data (Big Data)*, (pp. 5507-5512).

Guha, B., & Bandyopadhyay, G. (2016, March). Gold Price Forecasting Using ARIMA Model. *Journal of Advanced Management Science, 4*(2), 117-121.

Jackson, H., Muhammad, I. M., & Salaser , B. (2021). Time-Series Techniques for Prediction Sales in Cashback Websites. *iKSP Journal of Computer Science and Engineering*(1), 49-63.

Jain, G., & Prasad, R. R. (2020). Machine learning, Prophet and XGBoost algorithm: Analysis of Traffic Forecasting in Telecom Networks with time series data. *2020 8th International Conference on*

Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO) (pp. 893-897). IEEE.

Karabiber, O. A., & Xydis, G. (2019, March 10). Electricity Price Forecasting in the Danish Day-Ahead Market Using the TBATS, ANN and ARIMA Methods. *Energies, 12*(5), 928.

KullayaSwamy, A., Sarojamma, B., & Kiran Kumar, P. (2018, February). Time Series Modeling for Online Banking Transactions in India. *IJCRT, VI*(1), 2320-2882.

Li., Y. j., Yang, Y., Zhu, K., & Zhnag, J. (n.d.). Clothing Sale Forecasting by a Composite GRU–Prophet Model With an Attention Mechanism. *IEEE Transactions on Industrial Informatics*.

Lingyu , Z., Wenjie , B., Wenyi , Q., Liheng , T., & Yunhai , W. (2021). Time series forecast of sales volume based on XGBoost. *Journal of Physics: Conference Series*(1873 012067).

Maleki, A., Nasseri, S., Aminabad, M. S., & Hadi, M. (2018). Comparison of ARIMA and NNAR Models for Forecasting Water Treatment Plant's Influent Characteristics. *KSCE Journal of Civil Engineering*(22), 3233–3245.

Newaz, M. (2008). Comparing The Performance of Time Series Models for forecasting Exchange Rate. *BRAC University Journal, V*(2), 55-65.

P, H. K., & Patil, S. B. (2017). Forecasting Volatility with LSTM Techniques. *International Journal of Science and Research (IJSR)*, 2319-7064.

Sharma, S., & Bhatia, M. (2012). SHORT-TERM TRANSACTIONS FORECASTING USING TIME SERIES ANALYSIS: A CASE STUDY FOR INDIA. *Advances in Information Mining, 4*(1), 52-56.

Sudhakar, R., & Mahadevan, B. (Spring 2019). Nowcasting economic activity in India using payment systems data. *Journal of Payment Strategy & Systems, 13*(1), 72-81(10).

Tamatta, R. K. (2018). *Time series forecasting of hospital Inpatients and Day case waiting list using ARIMA, TBATS and Neural Network Models.* National College of Ireland, School of Computing.

Wu, C.-H., Ma, Y.-F., Lu, C.-C., & Lu, R.-S. (2018). A New Forecasting Framework for Bitcoin Price with LSTM. *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*, (pp. 168-175).

Yenidoğan, I., Çayir, A., Kozan, O., Dağ, T., & Arslan, Ç. (2018). Bitcoin Forecasting Using ARIMA and PROPHET. *2018 3rd International Conference on Computer Science and Engineering (UBMK)*, (pp. 621-624).

Aprigliano, V., Ardizzi, G., & Monteforte, L. (2017, April 6). Using the Payment System Data to Forecast the Italian GDP. *Bank of Italy Temi di Discussione (Working Paper) No. 1098*.

Aprigliano, V., Ardizzi, G., & Monteforte, L. (2019, October). Using Payment System Data to Forecast Economic Activity. *International Journal of Central Banking, 15*(4), 55-80.

Babii, A., Ghysels, E., & Striaukas, J. (2020, December 15). Machine Learning Time Series Regressions With an Application to Nowcasting.

Bentsen, K. N., & Gorea, D. (2021, April 22). Nowcasting and forecasting economic activity in Denmark using payment system data. *Working Paper - Danmarks National Bank*(177).

Buono, D., Mazzi, G. L., Kapetanios, G., Marcellino, M., & Papailias, F. (2017). Big data types for macroeconomic nowcasting. *Eurostat Review on National Accounts and Macroeconomic Indicators*, 94-145.

Chapman, J. T., & Desai, A. (2021, January 21). Using Payments Data to Nowcast Macroeconomic Variables During the Onset of COVID-19. *Bank of Canada Staff Working Paper*.

Dias , L. T., & Dias, A. C. (2017). Money talks – Nowcasting real economic activity with payment systems data. *Regional Statistics Conference.* Banco de Portugal, Lisbon, Portugal: International Statistical Institute.

Galbraith, J. W., & Tkacz, G. (2013). Nowcasting GDP: Electronic Payments, Data Vintages and the Timing of Data Releases. *Scientific Series*.

Galbraith, J. W., & Tkacz, G. (2018, April-June). Nowcasting with payments system data. *International Journal of Forecasting, 34*(2), 366-376.

Giannone, D., Reichlin, L., & Small, D. (2005). Nowcasting GDP and Inflation: The Real Time Informational Content of Macroeconomic Data Releases. *CEPR Discussion Papers 5178*.

León, C., & Ortega, F. (2018). Nowcasting economic activity with electronic payments data: A predictive modeling approach. *Banco De La Republica - Colombia*(1037).

Raj, J., Bhattacharyya, I., Behera, S. R., John, J., & Talwar, B. A. (2020). Modelling and Forecasting Currency Demand in India: A Heterodox Approach. *Reserve Bank of India Occasional Papers, 41*(1).

Raju, S., & Balakrishnan, M. (Spring 2019). Nowcasting economic activity in India using payment systems data. *Journal of Payments Strategy & Systems, 13*(1), 72-81.

Rooj, D., & Sengupta, R. (2021). Forecasting Private Consumption with Digital Payment Data: A Mixed Frequency Analysis. *ADBI Working Paper 1249*.
Bholat, D. (2015, January). Big Data and Central Banks. *Big Data & Society*, 86-93.

Gambacorta, L., Lombardi, M. J., Gorjon, S., & Conesa, C. (2015). The use of payment systems data as early indicators of economic activity. *Applied Economics Letters, 22:8*, 646-650.

Hall, A. S. (2018, November 13). Machine Learning Approaches to Macroeconomic Forecasting. *FEDERAL RESERVE BANK OF KANSAS CITY*, 63-81.

Hinge, D. (2019). 30 years of forecasting: have central banks learned anything?

Kim, S. (2020, September 2). Macroeconomic and Financial Market Analyses and Predictions Through Deep Learning. *Bank of Korea WP 2020-18*.

Microwave Consulting. (2021, March). *Analysis of India's payment system indicators in 2020.* Retrieved from Microsave: https://www.microsave.net/wp-content/uploads/2021/03/Analysis-of-India%E2%80%99s-payment-system-indicators-in-2020.pdf

Popowicz, J. E. (2021, February 11). Central banks look to shockproof models in wake of COVID-19. *Central Banking*.

Rijanto, E. (2018). "Building Pathways for Policy Making with Big Data. *BI – IFC / BIS INTERNATIONAL SEMINAR ON BIG DATA.* Bali.

Sivakumar, G. (2018). *Big Data for Central Banking.* Retrieved from IIT Bombay: https://www.cse.iitb.ac.in/~siva/talks/bdcentral.pdf

Wibisono, O., Ari, H. D., Widjanarti, A., Zulen, A. A., & Tissot, B. (2019, July 23-26). The use of big data analytics and artificial intelligence in central banking. *Irving Fisher Committee on Central Bank Statistics.* Bali: Irving Fisher Committee on Central Bank Statistics.

https://kevinkotze.github.io/ts-10-cointegration/