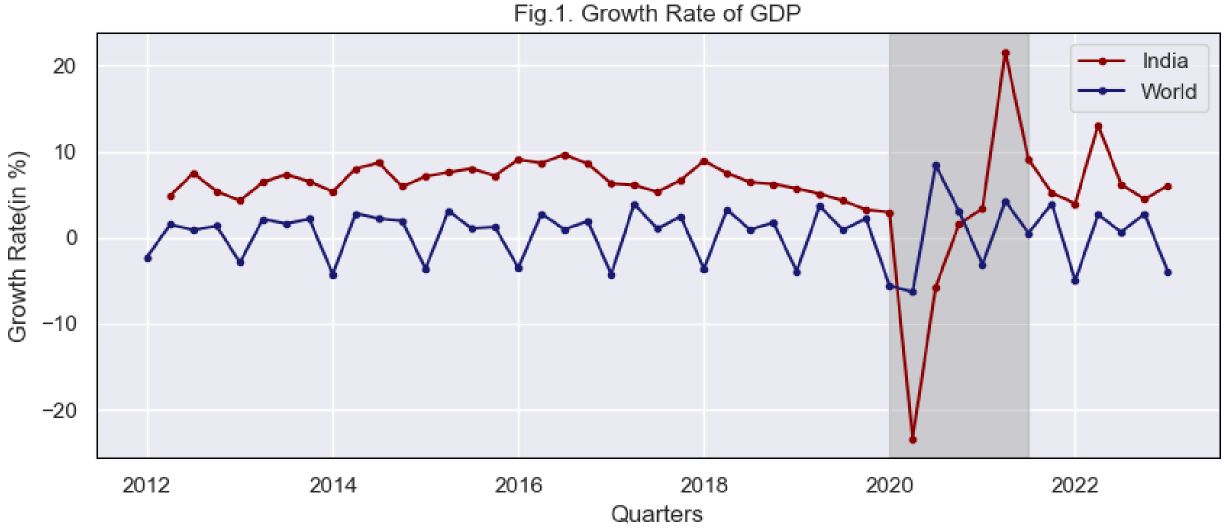
INDIA'S GDP PERFORMANCE SINCE 2012

Emerging as one of the fastest-growing economies, India's GDP grew by an average of 5.83% from Q1 2012 to Q4 2022, against the world benchmark of 0.48%. During the COVID-19 pandemic, output fell sharply by 23% in Q1 2020 before recovering relatively fast (Fig.1) with an average growth of 6.48% in the next 8 quarters. A sector-wise analysis reveals that financial, trade, manufacturing and agriculture sectors have contributed the highest to gross value added over 2012-2022.

Over the decade, the financial, trade and manufacturing sectors have performed significantly better with the robust growth rates (of 7.2% ,7.3% and 6.2% respectively). The government has put a lot of emphasis on the industrial sector through schemes like Productivity Linked Incentives, Ease of Doing Business, Make in India

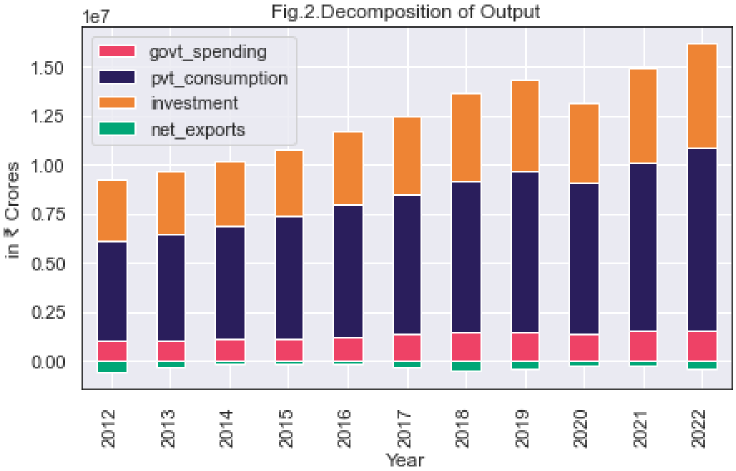
Fig.1. Growth Rate of GDP



initiative, etc. which explain the upward trend. The service sector has also maintained a stable trajectory of growth becoming a major driver of output. While agriculture falls behind with 3.8% growth rate, it is still credible for a developing country in the backdrop of severe climate change concerns.

Private consumption and investment have consistently remained the major drivers of economic growth for India (Fig.2). The economic growth post-COVID has been mostly consumption-led staying between 55% to 63% of GDP, exhibiting strong seasonality and volatility. Share of investment has also been a little volatile, falling to 24% at the wake of the pandemic Q1 2020 due to heightened uncertainty and worsening business confidence, later rising up to 35% in Q4 2022. India has become more integrated with the global economy in the past decades, which can be seen through the rise of volume of trade. Net exports of goods and services, however, has been stagnant.

Fig.2.Decomposition of Output



Forecasting Framework Of GDP

Our dataset contains quarterly time-series of GDP (at constant prices), Private Final Consumption Expenditure, Gross Fixed Capital Formation, Government Final Spending and Net Exports of goods and services, collected from MoSPI.

Objective: Forecast economic output of next 4 quarters using univariate and multivariate methods

Initial exploration of the GDP series through visualisation and Augmented-Dickey Fuller (ADF) Test reveals trend and seasonality components, making it non-stationary. ACF and PACF plots show a seasonal pattern every 4 quarters, which is why we will fit a seasonal ARIMA model.

SARIMA (p,d,q)(P,D,Q)S Model

Using the Box-Jenkins method, identification of optimum lags for seasonal orders (P,D,Q) and non-seasonal orders (p,d,q) is done first. To make the series stationary, there is a need to first difference it (d=1) and also seasonally difference it (D=1). Using Python, we automated models with different lag orders (p, q, P, Q) and chose the one with the least AIC metric.

Best Model: **SARIMA (1,1,1) (0,1,0)**

Estimation using statsmodels module gives the following results:

	coef	std err	z	P>|z|
ar.L1	0.8212	0.159	5.169	0.000
ma.L1	-0.9777	0.170	-5.741	0.000
sigma2	4.115e+10	4.39e-12	9.37e+21	0.000

The autoregressive and moving average coefficients are statistically significant. So, the GDP of quarter t is indeed influenced by the GDP of quarter (t-1) and past shock e_{t-1}.

Model evaluation

The residual diagnostic checks show that the residuals are uncorrelated (Ljung-Box Test and ACF plot of residuals), but they do not follow a normal distribution N~ (0,1). A possible reason for this could be the extreme values around Q1-2020 during the pandemic that skews the distribution.

Forecasting

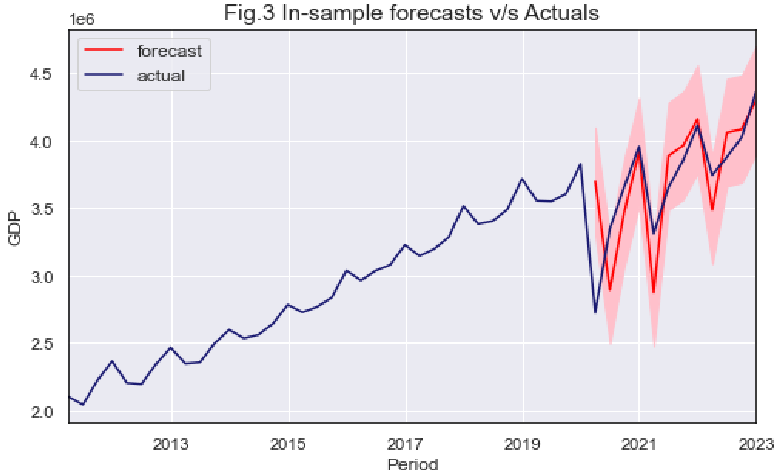
In-sample: To examine model accuracy, we predict the values for last 12 steps and plot them against actual values (Fig.3).

To measure how far the predictions are from real values:

Mean Absolute Error: 253587.26

Root Mean Squared Error: 360212.77

Fig.3 In-sample forecasts v/s Actuals



Out-of-sample Forecasting: Forecasting the GDP for future 4 quarters are given below.

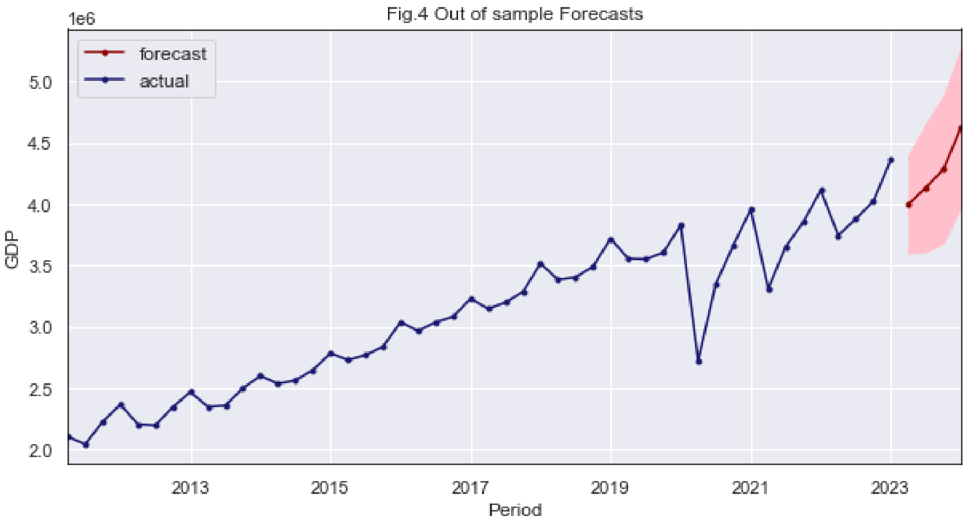


Table:1

SARIMA (1,1,1) (0,1,0)	
Predicted_GDP (in ₹ Crores)	
Q1 2023	3.997217e+06
Q2 2023	4.134034e+06
Q3 2023	4.281206e+06
Q4 2023	4.622187e+06

VAR (p) Model

Based on the Keynesian demand theory, economic output is determined by traditional drivers of aggregate demand which are- private consumption, government spending, investment and trade balance (net exports). For our purpose of forecasting GDP, we will use these real variables.

For the multivariate model, a vector autoregressive model is chosen since it is flexible and captures the interrelationships between multiple time-series variables better, by allowing feedback to occur within variables. Each variable has their own equation containing its own lagged values and lagged values of all other variables.

Checking stationarity
Upon visualisation and ADF test checks, it is seen that all the variables except net exports are trend non-stationary. So, we first difference them to make them stationary ~I (1).

Granger Causality Test
Before constructing a framework, it is essential to check whether variable X has any real influence that helps predict Y. Accounting for every variable, we can reject the null hypothesis of the Granger Causality Test and safely conclude that all the variables in the system are interchangeably causing each other.

Cointegration Test
Implementing the Johanssen cointegration test to examine whether there are long-run statistically significant relationships among the variables. Most variables exhibit cointegration.

Selecting Lag order
The model with 4 lags yields the minimum AIC and BIC. After fitting the data into a VAR (4), Most coefficients are not statistically significant and the error estimates seem to be higher than that of the ARIMA model. Iterating over lag orders= 1,2,3 it is found that VAR (1) produces the best results with least forecast errors of GDP.
After the estimation, in-sample predictions and forecasts for the next 4 quarters are done (Table.2). Durbin-Watson Statistics of every variable is close to 2, signifying no serial correlation among residuals. Additionally Impulse Response Functions (IRFs) and Forecast Error Variance Decomposition (FEVD) is analysed.

Best Model: **VAR (1)**
MAE: 199814.39
RSME: 252093.21

Impulse Response Functions (IRFs) trace the dynamic impact of a one-unit standard deviation "shock" to other variables in the system. In other words, how does the short-run dynamics look in response to a macroeconomic change like policy reforms?

From our IRFs,it is inferred that a shock in private consumption, GDP itself and investment initially increases the GDP in the first lag, followed by a steep decline in the next quarter and by the 5th quarter, the effect of shock dies down. Changes in net exports causes a similar but weaker oscillating effect on the GDP, but with higher uncertainty. Impact of changes in government spending on the GDP is less pronounced. The impact of an exogenous change in any of the variable converges back to its original value in around 6 quarters. So, there is no permanent structural change.

Forecast Error Variance Decomposition implies that initially, 56% of the variation in GDP is explained by its previous lag value while 20% is due to private consumption and 15% due to government spending. Investment only has around 4-5% contribution to the forecast variation while net exports has the lowest 3-4% influence.

Fig: **Impulse Responses**

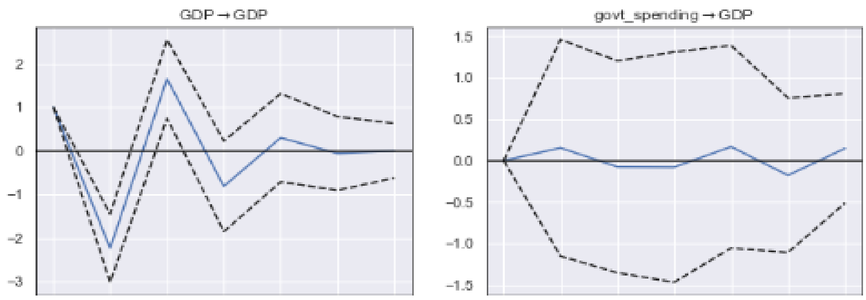


Table:2

VAR (1)	
Predicted_GDP (in ₹ Crores)	
Q1 2023	3.942597e+06
Q2 2023	4.356879e+06
Q3 2023	4.147819e+06
Q4 2023	4.347823e+06

Which model is better?

The multivariate model VAR (1) performs better in forecasting economic output than the univariate SARIMA model in terms of accuracy metrics like Mean Absolute Error and Mean Squared Error. From both the models, it is quite certain that GDP is the most affected by its own past values than any other factors. However, the VAR captures the complex feedback effects of other real variables like private consumption, government spending and investment on the economic output, accounting For the rest of the variation. The VAR method also enables additional analysis of exogenous changes in the system, which is not available in the univariate part.