

Forecasting Inflation in Pakistan

Using Multiple Linear Regression Models

Group Members:

Muhammad Abdullah Aamir (23i-2538)

Muhammad Abdullah Ali (23i-2523)

Submission Date: 4 May 2025

Department of Statistics

Your University Name

Contents

1	Introduction	2
1.1	The Significance of Inflation in Pakistan	2
1.2	Project Objective	2
1.3	Methodology and Modeling Approach	2
2	Literature Review	3
2.1	Introduction	3
2.2	Overview of Inflation Dynamics in Pakistan	3
2.3	Analysis of Key Forecasting Models	4
2.3.1	Model 1: Autoregressive Distributed Lag (ARDL) Models	4
2.3.2	Model 2: Vector Autoregressive (VAR) Models	4
2.3.3	Model 3: Leading Indicators Models (LIM)	4
2.3.4	Model 4: Autoregressive Integrated Moving Average (ARIMA) Models	4
2.4	Comparative Assessment and Synthesis	4
2.5	Broader Context and Recent Trends	5
2.6	Conclusion	5
3	Description of Variables and Data Description	5
3.1	Dependent Variable	6
3.2	Independent Variables	6
3.3	Data Sources and Time Period	6
3.4	Data Preprocessing	7
4	Estimation of the Models	7
4.1	ARIMA Model Estimation	7
4.1.1	Selected ARIMA Models	8
4.2	Ridge Regression Estimation	9
4.3	LASSO Regression Estimation	9
4.4	Elastic Net Regression Estimation	9
5	Results and Conclusion	10
5.1	ARIMA Model Results and Performance	10
5.1.1	Model Fit and Residual Diagnostics (Training Data)	10
5.1.2	Forecasting Performance Evaluation (Test Set)	11
5.2	Ridge Regression Results and Performance	11
5.3	LASSO Regression Results and Performance	12
5.4	Elastic Net Regression Results and Performance	13
5.5	Comparison of Model Performance and Overall Conclusion	14
5.5.1	Comparative Summary of Model Performance	14
5.5.2	Overall Conclusion	14
5.5.3	Recommendations for Future Work	16

1 Introduction

Inflation, defined as the sustained increase in the general price level of goods and services in an economy over a period of time, is a critical macroeconomic indicator with profound implications for economic stability, growth, and the welfare of citizens. High and volatile inflation erodes purchasing power, distorts investment decisions, increases uncertainty, and can lead to social and political instability. For a developing economy like Pakistan, which has historically faced challenges related to price stability, understanding and forecasting inflation is paramount for effective economic policymaking. Accurate inflation forecasts enable policymakers, businesses, and individuals to make informed decisions, facilitating better resource allocation, monetary policy formulation, and financial planning.

1.1 The Significance of Inflation in Pakistan

Pakistan's economy has frequently grappled with inflationary pressures stemming from a variety of factors, including supply chain disruptions, exchange rate fluctuations, fiscal imbalances, global commodity price movements, and structural issues. The persistent challenge of managing inflation has significant consequences, impacting the cost of living, particularly for low-income households, affecting the competitiveness of exports, and influencing the overall investment climate. Therefore, developing robust models to forecast inflation is not merely an academic exercise but a practical necessity for achieving macroeconomic stability and fostering sustainable economic development in the country. This project focuses specifically on the context of Pakistan, aiming to contribute to a better understanding of its inflationary dynamics.

1.2 Project Objective

The primary objective of this project is to develop and evaluate quantitative models for forecasting inflation in Pakistan. By identifying key macroeconomic variables that influence inflation and employing appropriate statistical techniques, we aim to build models that can provide reliable predictions of future price movements. This will involve collecting relevant data, performing exploratory data analysis, constructing forecasting models, and evaluating their performance.

1.3 Methodology and Modeling Approach

To achieve the project objective, we will employ various regression techniques, which are fundamental statistical methods used to model the relationship between a dependent variable and one or more independent variables. Regression analysis allows us to understand how changes in the independent variables are associated with changes in the dependent variable, and crucially, to make predictions about the dependent variable based on the values of the independent variables. In the context of inflation forecasting, regression helps us quantify the impact of potential drivers of inflation and project future inflation rates.

Specifically for forecasting inflation in Pakistan, this project will explore and compare the performance of several regression-based models:

- **ARIMA (AutoRegressive Integrated Moving Average):** A time series forecasting model that models the next step in a time series based on linear combinations of past observations and past errors.
- **Ridge Regression:** A regularization technique used in linear regression that adds a penalty term to the loss function to shrink the regression coefficients towards zero, which can help prevent overfitting, especially when dealing with multicollinearity.
- **LASSO (Least Absolute Shrinkage and Selection Operator) Regression:** Another regularization technique that adds a penalty term based on the absolute values of the coefficients. LASSO can perform variable selection by shrinking the coefficients of less important variables to exactly zero.
- **Elastic Net Regression:** A hybrid regularization method that combines the penalties of both Ridge and LASSO regression, offering a balance between shrinking coefficients and performing variable selection.

The project will involve the following key steps:

- Data collection from credible sources for the chosen variables.
- Analysis of summary statistics and visualization of the data to understand the characteristics and relationships between variables.
- Construction and estimation of the specified regression models.
- Evaluation of model assumptions and statistical significance of the independent variables.
- Comparison of model predictions with actual inflation data to assess forecasting accuracy using metrics like Mean Squared Error (MSE) and graphical comparisons.

2 Literature Review

2.1 Introduction

Inflation, a sustained increase in the general price level, significantly impacts developing nations like Pakistan by eroding purchasing power and affecting economic stability. Pakistan has historically experienced volatile inflation, highlighting the critical need for accurate forecasting to inform policy and foster sustainable growth. This review analyzes existing research on inflation in Pakistan, focusing on four key models relevant to this project: ARDL, VAR, Leading Indicators, and ARIMA. We will examine their theoretical bases, empirical findings, and limitations to understand their strengths and weaknesses in the Pakistani context.

2.2 Overview of Inflation Dynamics in Pakistan

Pakistan's inflation history shows significant fluctuations, including a prolonged period of double-digit inflation (FY2008-FY2012) and a recent surge peaking in FY2023 before moderating in 2024. Key drivers identified in the literature include monetary factors (broad money and private credit growth), fiscal policy (government spending and deficits),

exchange rate depreciation (impacting import costs), global commodity prices (especially oil and food), supply-side issues (crop yields), and political/economic uncertainty. The complex interplay of these factors necessitates comprehensive forecasting models.

2.3 Analysis of Key Forecasting Models

This section details four prominent models used for inflation forecasting in Pakistan:

2.3.1 Model 1: Autoregressive Distributed Lag (ARDL) Models

ARDL models analyze dynamic relationships, capturing both short-run and long-run effects of variables like output gap, oil prices, interest rates, and money supply. Studies show ARDL models can be effective, particularly in high inflation regimes and for longer-term forecasts in low inflation regimes, sometimes outperforming benchmarks. However, appropriate lag selection and potential misspecification are considerations.

2.3.2 Model 2: Vector Autoregressive (VAR) Models

VAR models treat all variables as endogenous, modeling their interrelationships (e.g., between inflation, monetary aggregates, credit, and exchange rates). While useful for understanding dynamics, some research suggests VAR models may not consistently outperform simpler ARIMA models in forecasting accuracy for Pakistan. Limitations include sensitivity to lag selection and potential over-parameterization.

2.3.3 Model 3: Leading Indicators Models (LIM)

LIM identify variables that precede inflation changes, such as broad money growth and private sector credit growth. Studies report good ex-post forecast quality for LIM in Pakistan, suggesting their suitability for short-term predictions. A limitation is their empirical basis, which may require periodic re-evaluation of indicators.

2.3.4 Model 4: Autoregressive Integrated Moving Average (ARIMA) Models

ARIMA models are univariate time series models forecasting based on past values and errors. Widely used for Pakistan's CPI, they are statistically sound but do not directly incorporate external macroeconomic variables. They may also struggle with predicting turning points or external shocks. Despite limitations, ARIMA models serve as a fundamental tool and benchmark.

2.4 Comparative Assessment and Synthesis

Table 1 provides a comparative summary of the four models.

No single model consistently outperforms others across all conditions in Pakistan. Their effectiveness depends on economic context and forecast horizon. ARDL is strong in high inflation, LIM is for the short term, and ARIMA is a solid baseline.

Table 1: Comparison of Inflation Forecasting Models for Pakistan

Model	Theoretical Basis	Key Variables	Forecasting Performance (General Findings)	Strengths	Limitations
ARDL	Statistical model capturing short-run/long-run relationships	Output gap, oil prices, interest rates, money supply, world CPI, industrial production, exchange rate	Effective, esp. for high inflation regimes	Handles mixed integration orders; captures dynamics.	Lag selection; misspecification.
VAR	Statistical model (all variables endogenous)	Discount rate, T-bill rate, credit growth, monetary aggregates, output gap, exchange rate	May not outperform simpler models; illustrates monetary relationships.	Captures multivariate interactions.	Lag sensitivity; over-parameterization; complex interpretation.
Leading Indicators	Empirical (variables preceding inflation)	Broad money growth, private credit growth	Good ex-post quality; suitable for short-term.	Uses practical predictive variables.	Empirical basis; indicators may change.
ARIMA	Statistical (past time series values)	Past inflation values (CPI)	Performance varies; can be effective with GARCH.	Statistically rigorous; widely understood.	Univariate; struggles with turning points/shocks; requires stationarity.

2.5 Broader Context and Recent Trends

Beyond these four, other models like Bayesian VAR, Dynamic Factor Models, and Phillips Curve models are used. Recent trends show increasing interest in Machine Learning techniques (ANN, Ridge, LASSO, Elastic Net), which have demonstrated potential for higher accuracy, particularly in data-rich environments. There’s also a focus on high-frequency data and alternative sources for timely forecasting.

2.6 Conclusion

This review analyzed ARDL, VAR, LIM, and ARIMA models for Pakistan’s inflation forecasting, highlighting their characteristics and performance. While each has merits, no single best model exists universally. Future research should focus on granular analysis of specific models, rigorous comparison of traditional vs. ML techniques, exploring hybrid models, improving accuracy during volatility, incorporating inflation expectations, and utilizing high-frequency data.

3 Description of Variables and Data Description

This section provides a detailed description of the variables used in this study for forecasting inflation in Pakistan, their definitions, sources, the time period covered by the data, and the preprocessing steps undertaken.

3.1 Dependent Variable

The dependent variable in this study is ****Headline Consumer Price Index (CPI) Year-on-Year Inflation****. This is the most commonly used measure of overall inflation, reflecting the percentage change in the price of a basket of consumer goods and services compared to the same month in the previous year. This variable captures the primary phenomenon we aim to forecast.

3.2 Independent Variables

Based on the literature review and the requirements of the project to include at least five independent variables, a comprehensive set of macroeconomic indicators and their lags were considered as potential predictors of inflation. The selection of these variables is guided by economic theory and empirical findings from previous studies on inflation determinants in Pakistan. The independent variables broadly fall into the following categories:

- **Monetary Variables:** Including measures of money supply growth (e.g., Broad Money (M2) growth) and credit growth (e.g., Private Sector Credit growth), which capture demand-side pressures.
- **Fiscal Variables:** Such as government expenditure and revenue, reflecting the stance of fiscal policy and its potential impact on aggregate demand.
- **Exchange Rate:** The Pak Rupee / US Dollar exchange rate, which accounts for imported inflation through the pass-through effect.
- **Real Activity Measures:** Indicators like industrial production or GDP growth, representing the state of the real economy and potential Phillips-curve effects.
- **External Factors:** Primarily global commodity prices (e.g., oil price indices, food price indices), which significantly influence domestic prices in an import-dependent economy like Pakistan.
- **Lagged Variables:** Various lags of the dependent variable (inflation) and the independent variables were included to capture the time-lagged effects and dynamics inherent in macroeconomic relationships. The specific set of predictors used in the regression models, including their lags, totaled 60 variables as indicated in the model estimation output.

3.3 Data Sources and Time Period

The data for the dependent and independent variables were collected from authentic and publicly available sources, consistent with the project requirements. These sources include:

- **State Bank of Pakistan (SBP):** A primary source for monetary, credit, and exchange rate data.
- **Pakistan Bureau of Statistics (PBS):** The official source for Consumer Price Index (CPI) and Wholesale Price Index (WPI) data, as well as industrial production statistics.

- **International Financial Statistics (IFS) / World Development Indicators (WDI) (IMF/World Bank):** Providing international data, including global commodity prices and potentially cross-country comparisons if needed (though the focus is on Pakistan).

The dataset comprises monthly observations. Based on the training and test split sizes observed in the model outputs (e.g., 64/16 or 81/24 months), the total time period covered by the data is approximately 80 to 105 months, spanning roughly from late 2016/early 2017 up to early 2025. A precise start and end date will be determined based on the final compiled dataset.

3.4 Data Preprocessing

Prior to model estimation, the raw data underwent several preprocessing steps:

- **Handling Missing Values:** Missing observations were addressed, potentially through interpolation (e.g., linear interpolation using `na.approx` as seen in the ARIMA script output) or other appropriate methods. Any remaining missing values after interpolation were likely handled by omitting those time periods.
- **Creating Lagged Variables:** Relevant lagged values of both the dependent and independent variables were generated to capture dynamic relationships and fulfill the requirements for time series and regression models.
- **Scaling:** The independent variables and the dependent variable were scaled (standardized) by subtracting the mean and dividing by the standard deviation. This step is particularly important for regularization techniques like Ridge, LASSO, and Elastic Net, as it ensures that the penalty terms are applied fairly across all coefficients, regardless of the original scale of the variables. The model estimation was performed using the scaled data, and predictions were subsequently unscaled for evaluation on the original scale.
- **Train/Test Split:** The data was split into a training set (approximately 80% of the data) used for model estimation and tuning, and a test set (the remaining 20

These steps ensure that the data is in an appropriate format for the chosen forecasting models and that the evaluation provides a realistic assessment of their out-of-sample performance.

4 Estimation of the Models

This section details the specification and estimation of the forecasting models used in this project for Pakistan: ARIMA, Ridge Regression, LASSO Regression, and Elastic Net Regression.

4.1 ARIMA Model Estimation

This subsection focuses on the estimation of the ARIMA models. ARIMA (AutoRegressive Integrated Moving Average) models are a class of time series models that capture the temporal dependencies within a single series. An $ARIMA(p,d,q)(P,D,Q)[m]$ model

is characterized by non-seasonal parameters (p, d, q) and seasonal parameters (P, D, Q) with a seasonal period m . The parameters represent:

- p (AR order): The number of lag observations included in the model.
- d (Differencing order): The number of times the raw observations are differenced to make the time series stationary.
- q (MA order): The size of the moving average window, representing the number of lagged forecast errors.
- P (Seasonal AR order): The number of seasonal lag observations.
- D (Seasonal Differencing order): The number of times seasonal differencing is applied.
- Q (Seasonal MA order): The number of seasonal lagged forecast errors.
- m : The number of periods in each season (e.g., 12 for monthly data).

A drift term may also be included to account for a constant trend in the differenced series.

For each of the four inflation series, the `auto.arima()` function in R was employed to automatically select the optimal ARIMA model structure based on minimizing information criteria (such as AIC, AICc, or BIC) on the training data. The estimated coefficients for the selected models are presented in the results section.

4.1.1 Selected ARIMA Models

Based on the training data, the following ARIMA models were selected by the `auto.arima()` function for each inflation series:

CPI_YoY_National (National CPI, Year-on-Year) The selected model for National CPI Year-on-Year was `ARIMA(1,1,2)(1,0,1)[12] with drift`. This model includes one non-seasonal AR term, one order of non-seasonal differencing, two non-seasonal MA terms, one seasonal AR term, zero orders of seasonal differencing, one seasonal MA term with a seasonal period of 12, and a drift term.

CPI_YoY_NFNE_Urban (Urban Non-Food Non-Energy CPI, Year-on-Year) The selected model for Urban NFNE CPI Year-on-Year was `ARIMA(1,1,2)(2,0,0)[12] with drift`. This model includes one non-seasonal AR term, one order of non-seasonal differencing, two non-seasonal MA terms, two seasonal AR terms, zero orders of seasonal differencing, zero seasonal MA terms with a seasonal period of 12, and a drift term.

CPI_YoY_Trimmed_Urban (Urban Trimmed CPI, Year-on-Year) The selected model for Urban Trimmed CPI Year-on-Year was `ARIMA(1,1,0)(0,0,1)[12] with drift`. This is a simpler model including one non-seasonal AR term, one order of non-seasonal differencing, zero non-seasonal MA terms, zero seasonal AR terms, zero orders of seasonal differencing, one seasonal MA term with a seasonal period of 12, and a drift term.

WPI_YoY_General (WPI General, Year-on-Year) The selected model for WPI General Year-on-Year was `ARIMA(0,1,0)(1,0,1)[12] with drift`. This model includes zero non-seasonal AR terms, one order of non-seasonal differencing, zero non-seasonal MA terms, one seasonal AR term, zero orders of seasonal differencing, one seasonal MA term with a seasonal period of 12, and a drift term.

4.2 Ridge Regression Estimation

Ridge Regression is a regularization technique applied to linear regression that addresses multicollinearity and prevents overfitting by adding an L2 penalty term ($\lambda \sum \beta_j^2$) to the standard least squares objective function. This penalty shrinks the regression coefficients towards zero but does not set them exactly to zero, meaning all predictor variables are retained in the model.

The model was estimated using the scaled training data. The regularization parameter λ was tuned using 10-fold cross-validation. The optimal λ was selected using the one-standard-error rule (`lambda.1se`), resulting in $\lambda = 112.5481$. This relatively high value suggests that substantial regularization was applied, likely due to the presence of multicollinearity among the 60 predictor variables.

After tuning, the Ridge regression model was fitted to the entire scaled training dataset using the selected λ . All 60 predictor variables were retained with non-zero coefficients, although their magnitudes were shrunk by the regularization. Interpretation of the individual impact of each predictor is challenging due to the large number of variables and the nature of the Ridge penalty.

4.3 LASSO Regression Estimation

LASSO (Least Absolute Shrinkage and Selection Operator) Regression is another regularization technique for linear regression that adds an L1 penalty term ($\lambda \sum |\beta_j|$) to the objective function. A key feature of the LASSO penalty is its ability to perform automatic variable selection by shrinking the coefficients of less important predictors exactly to zero.

The LASSO model was estimated using the scaled training data. The regularization parameter λ was tuned using 10-fold cross-validation, and the optimal λ was selected using the one-standard-error rule (`lambda.1se`), resulting in $\lambda = 0.2854$.

The LASSO model achieved significant sparsity, effectively selecting a much smaller subset of predictors by setting the coefficients of many variables to zero. The variables selected by the LASSO model (those with non-zero coefficients) were: `GovRev_YoY`, `GovExp_YoY_L1`, `GovExp_YoY_L3`, `GovExp_YoY_L6`, `GovBorrow_YoY_L6`, and `Core_Trimmed`. This indicates that fiscal variables (government revenue, expenditure at various lags, and government borrowing at a lag) and the underlying trend captured by trimmed core inflation were identified as the most important predictors by the LASSO model within this dataset and timeframe.

4.4 Elastic Net Regression Estimation

Elastic Net Regression is a hybrid regularization method that combines the penalties of both Ridge (L2) and LASSO (L1) regression. It is controlled by two parameters: α and λ . The α parameter determines the mix between the L1 and L2 penalties ($\alpha = 1$ for pure

LASSO, $\alpha = 0$ for pure Ridge), and λ controls the overall strength of the regularization. Elastic Net offers a balance between shrinking coefficients (like Ridge) and performing variable selection (like LASSO), and it can be particularly useful when dealing with groups of highly correlated predictors.

The Elastic Net model was estimated using the scaled training data. Both α and λ were tuned using a grid search approach combined with 10-fold cross-validation. The optimal α was selected based on minimizing the cross-validation error, and the corresponding λ was chosen using the one-standard-error rule (`lambda.1se`). The tuning process selected $\alpha = 0.9$ and $\lambda = 0.3171$. The optimal α being close to 1 indicates that the model favors the L1 penalty, behaving very similarly to the LASSO model for this dataset.

The Elastic Net model, with $\alpha = 0.9$, also performed variable selection, retaining a small set of predictors. The variables selected by the Elastic Net model (those with non-zero coefficients) were: `GovRev_YoY`, `GovExp_YoY_L1`, `Core_Trimmed_L1`, `GovExp_YoY_L3`, `GovExp_YoY_L6`, `GovBorrow_YoY_L6`, and `Core_Trimmed`. This set of 7 selected variables is very similar to the set selected by LASSO, with the notable addition of the first lag of trimmed core inflation (`Core_Trimmed_L1`). This further reinforces the importance of fiscal factors and core inflation trends as identified by the penalized regression methods.

5 Results and Conclusion

This section presents the results of fitting the selected models and evaluates their forecasting performance on the held-out test set. This includes presenting model outputs, discussing statistical significance of variables (for regression models), analyzing residual diagnostics, evaluating forecast accuracy metrics, and comparing predicted values with actuals graphically and using MSE. The overall conclusion of the project will also be presented in this section.

5.1 ARIMA Model Results and Performance

This subsection presents the results and evaluates the performance of the estimated ARIMA models.

5.1.1 Model Fit and Residual Diagnostics (Training Data)

After fitting the selected ARIMA models to the training data, residual diagnostics were performed to assess if the models adequately captured the time series patterns, leaving only white noise in the residuals. The Ljung-Box test was used to test for autocorrelation in the residuals.

- **CPI_YoY_National:** Ljung-Box test p-value = 0.3737. Since the p-value is greater than 0.05, we fail to reject the null hypothesis of no autocorrelation in the residuals. This suggests the model fits the training data well.
- **CPI_YoY_NFNE_Urban:** Ljung-Box test p-value = 0.9997. The high p-value indicates no significant autocorrelation in the residuals, suggesting a good fit to the training data.
- **CPI_YoY_Trimmed_Urban:** Ljung-Box test p-value = 0.9993. The high p-value indicates no significant autocorrelation in the residuals, suggesting a good fit to the training data.

- **WPI_YoY_General:** Ljung-Box test p-value = 0.9901. The high p-value indicates no significant autocorrelation in the residuals, suggesting a good fit to the training data.

For all four series, the residual diagnostics indicate that the selected ARIMA models provide a good fit to the historical patterns in the training data, as evidenced by the lack of significant autocorrelation in the residuals. Residual diagnostic plots were also generated (e.g., `arima_residuals_*.pdf`) to visually inspect the residuals.

5.1.2 Forecasting Performance Evaluation (Test Set)

The performance of the fitted ARIMA models was evaluated on the test set using standard forecasting accuracy metrics. The table below summarizes the key metrics for the test set forecasts:

Discussion of Forecast Accuracy:

The forecasting performance on the test set reveals significant challenges for the ARIMA models during this specific period.

- For **National CPI**, **Urban NFNE CPI**, and **WPI General**, the error metrics (RMSE, MAE, MAPE) are very high, indicating large discrepancies between the forecasted and actual values. The Mean Error (ME) being negative for these series suggests a tendency for the models to over-forecast inflation during the test period. Crucially, the Theil's U statistic is significantly greater than 1 for all three series (42.48, 8.68, and 24.34 respectively). A Theil's U greater than 1 implies that the ARIMA model's forecast is worse than a simple naive forecast (using the last observed value as the prediction), highlighting poor performance relative to a basic benchmark.
- The **Urban Trimmed CPI** model showed relatively better performance compared to the other three series, with lower RMSE (4.32), MAE (3.87), and MAPE (27.33%). However, its Theil's U (4.435) is still greater than 1, indicating that even for this series, the ARIMA forecast was worse than a naive forecast on the test set. The positive ME (1.54) suggests a slight tendency to under-forecast for this series.

Forecast plots comparing the actual values to the model forecasts on the test set were generated (e.g., `arima_forecast_vs_actual_*.pdf`). These plots visually confirm the observed inaccuracies, showing that the ARIMA models struggled to capture the magnitude and turning points of inflation during the test period, particularly the sharp decline observed in some series.

The poor performance of the univariate ARIMA models on the test set, despite fitting the training data well, suggests that factors beyond the historical pattern of the individual inflation series were significant drivers of inflation during the test period. This underscores the potential need for multivariate models that incorporate external economic variables, which aligns with the approach of using Ridge, LASSO, and Elastic Net regression.

5.2 Ridge Regression Results and Performance

This subsection presents the results and evaluates the performance of the estimated Ridge Regression model.

Model Fit (Training Data): The in-sample R-squared on the scaled training data was 0.7809. This indicates a moderate fit to the training data, meaning the model explains approximately 78.09% of the variance in the scaled Headline CPI within the training set.

Forecasting Performance Evaluation (Scaled Test Set): The performance of the Ridge Regression model was evaluated on the scaled test set. The key accuracy metrics are:

- **RMSE (Scaled):** 4.7914
- **MAE (Scaled):** 4.0006

These metrics represent the average magnitude of the errors on the scaled test data.

Forecasting Performance Evaluation (Unscaled Test Set) - Crucial Caveat: The reported performance metrics on the original (unscaled) test set (RMSE: 128.04, MAE: 123.21, MAPE: 1447.29%) are extremely high and appear unreliable. As noted in the analysis output, there seems to be an issue with the unscaling process for the regression model predictions. Therefore, the unscaled metrics should be interpreted with extreme caution and are not used for reliable comparison with the actual unscaled inflation values or the ARIMA model's unscaled performance. Comparisons of Ridge performance with other regression models will rely primarily on the metrics calculated on the scaled test data.

Discussion of Results: Ridge regression served as a baseline regularized model. Its performance on the scaled test set, with an RMSE of 4.7914 and MAE of 4.0006, was significantly worse compared to the LASSO and Elastic Net models (as will be shown in the comparative summary). The Ridge model retained all 60 predictors, making it less interpretable than models that perform variable selection. The high optimal lambda value chosen during tuning suggests that multicollinearity might be a significant issue in the predictor set, which Ridge attempts to mitigate by shrinking coefficients. However, this shrinkage alone did not lead to superior out-of-sample performance in this case.

5.3 LASSO Regression Results and Performance

This subsection presents the results and evaluates the performance of the estimated LASSO Regression model.

Model Fit (Training Data): The in-sample R-squared on the scaled training data was 0.9907. This indicates an extremely high fit to the training data, explaining over 99% of the variance in the scaled Headline CPI within the training set. While a high R-squared is generally desirable, such a high value, even with cross-validation, warrants caution regarding potential overfitting to the training data.

Variable Selection: A key result of the LASSO model is the automatic variable selection. The LASSO model, with $\lambda = 0.2854$, selected the following 6 predictors with non-zero coefficients:

- GovRev_YoY
- GovExp_YoY_L1
- GovExp_YoY_L3
- GovExp_YoY_L6

- GovBorrow_YoY_L6
- Core_Trimmed

This result highlights the importance of fiscal variables (government revenue, government expenditure at lags 1, 3, and 6, and government borrowing at a lag) and the current value of trimmed core inflation as the most influential predictors for Headline CPI inflation according to the LASSO model.

Forecasting Performance Evaluation (Scaled Test Set): The performance of the LASSO Regression model was evaluated on the scaled test set. The key accuracy metrics are:

- RMSE (Scaled): 2.6824
- MAE (Scaled): 2.3821

These metrics indicate the average magnitude of the errors on the scaled test data.

Forecasting Performance Evaluation (Unscaled Test Set) - Crucial Caveat: Similar to the Ridge model, the reported performance metrics on the original (unscaled) test set (RMSE: 128.01, MAE: 110.93, MAPE: 1036.76%) are extremely high and unreliable due to the apparent issue with the unscaling process. These unscaled metrics should not be used for reliable interpretation or comparison on the original inflation scale.

Discussion of Results: The LASSO regression model demonstrated the best predictive performance among all evaluated models based on the scaled test data metrics (lowest RMSE and MAE). Its ability to select a small, interpretable set of variables from the large pool of 60 predictors is a significant advantage. The selected variables (fiscal factors and core inflation) align reasonably with economic intuition regarding drivers of inflation in Pakistan. However, the extremely high in-sample R-squared warrants careful consideration of potential overfitting, even though cross-validation was used for tuning. The variable selection property of LASSO makes the model more parsimonious and potentially more robust and interpretable than Ridge regression.

5.4 Elastic Net Regression Results and Performance

This subsection presents the results and evaluates the performance of the estimated Elastic Net Regression model.

Model Fit (Training Data): The in-sample R-squared on the scaled training data was 0.9906. Similar to the LASSO model, this indicates an extremely high fit to the training data, explaining over 99% of the variance in the scaled Headline CPI within the training set. This also raises potential concerns about overfitting, similar to the LASSO model.

Variable Selection: The Elastic Net model, with optimal parameters $\alpha = 0.9$ and $\lambda = 0.3171$, also performed variable selection. The variables selected by the Elastic Net model (those with non-zero coefficients) were:

- GovRev_YoY
- GovExp_YoY_L1
- Core_Trimmed_L1
- GovExp_YoY_L3

- GovExp_YoY_L6
- GovBorrow_YoY_L6
- Core_Trimmed

This set of 7 selected variables is very similar to the set selected by LASSO, with the notable addition of the first lag of trimmed core inflation (**Core_Trimmed_L1**). The high optimal α value (0.9) confirms that the model leaned heavily towards the L1 penalty, resulting in a sparse solution similar to LASSO.

Forecasting Performance Evaluation (Scaled Test Set): The performance of the Elastic Net Regression model was evaluated on the scaled test set. The key accuracy metrics are:

- **RMSE (Scaled):** 2.7774
- **MAE (Scaled):** 2.4324

These metrics indicate the average magnitude of the errors on the scaled test data.

Forecasting Performance Evaluation (Unscaled Test Set) - Crucial Caveat: As with the other regression models, the reported performance metrics on the original (unscaled) test set (RMSE: 126.55, MAE: 108.68, MAPE: 991.14%) are extremely high and unreliable due to the apparent issue with the unscaling process. These unscaled metrics should not be used for reliable interpretation or comparison on the original inflation scale.

Discussion of Results: The Elastic Net model, with an optimal α close to 1, performed very similarly to the LASSO model in terms of variable selection and scaled test set performance. Its scaled RMSE (2.7774) and MAE (2.4324) were slightly higher than LASSO but significantly better than Ridge. The selected variables reinforce the findings from LASSO regarding the importance of fiscal factors and core inflation. Elastic Net provides a robust alternative to LASSO, particularly in scenarios with highly correlated predictors, although the high α here suggests that the L1 penalty's variable selection capability was the dominant factor for this dataset. Similar concerns about potential overfitting due to the high in-sample R-squared apply.

5.5 Comparison of Model Performance and Overall Conclusion

This subsection provides a comparative summary of the performance of all evaluated models and presents the overall conclusions of the project.

5.5.1 Comparative Summary of Model Performance

The table below summarizes the key characteristics and performance metrics for the four models evaluated for forecasting Pakistan's Headline CPI.

5.5.2 Overall Conclusion

Based on the analysis and evaluation of the four forecasting models for Pakistan's Headline CPI inflation, the following conclusions can be drawn:

Table 2: Comparative Summary of Inflation Forecasting Model Performance for Pakistan

Model	Approach	Variable Selection	Key Strength	Key Weakness	Scaled Test RMSE	Scaled Test MAE	MASE (AR)
ARIMA	Time Series	No (Univariate)	Standard benchmark; Captures time dynamics	Poor Accuracy ($MASE_{t 1}$); Ignores external factors	N/A	N/A	2.36
Ridge	Regularization	No (Shrinkage)	Handles multicollinearity	No variable selection; Poorer accuracy	4.7914	4.0006	N/A
LASSO	Regularization	Yes	Best Accuracy (Scaled); Sparsity; Interpretability	Potential instability with correlated predictors; Overfitting risk	2.6824	2.3821	N/A
Elastic Net	Regularization	Yes	Good accuracy; Handles correlated groups well	Slightly less accurate than Lasso here; Overfitting risk	2.7774	2.4324	N/A

- Superiority of Variable Selection Models:** For forecasting Headline CPI using the provided extensive set of 60 potential predictors, the models employing variable selection (LASSO and Elastic Net) significantly outperformed Ridge Regression based on the scaled test set metrics (RMSE and MAE). This highlights the benefit of identifying and focusing on a smaller, more relevant subset of key drivers in a high-dimensional predictive modeling context.
- Inadequacy of Univariate ARIMA:** The standard univariate ARIMA model, while fitting the training data well based on residual diagnostics, demonstrated poor out-of-sample forecasting ability for National CPI inflation during the test period. Its MASE value significantly greater than 1 indicates that its forecasts were less accurate than a simple naive forecast (using the last observed value). This underscores the limitations of relying solely on the past history of inflation and suggests that external factors, policy changes, and structural breaks prevalent in Pakistan's economy likely play a crucial role that univariate models cannot capture. The results for other ARIMA targets (NFNE Urban, Trimmed Urban, WPI) also showed MASE values greater than 1, indicating similar poor out-of-sample performance relative to a naive benchmark.
- Identified Key Drivers:** The penalized regression models (LASSO and Elastic Net) consistently pointed towards the importance of fiscal variables (government revenue, expenditure, and borrowing at various lags) and core inflation trends (Core_Trimmed and its lag for Elastic Net) as significant predictors within this dataset and timeframe. This aligns with economic intuition regarding demand-pull pressures and inflation persistence as drivers of inflation in Pakistan.
- Overfitting and Scaling Concerns:** The extremely high in-sample R-squared values observed for both LASSO and Elastic Net (above 0.99) warrant caution regarding potential overfitting to the training data, despite the use of cross-validation

for tuning. More significantly, the apparent failure of the prediction unscaling process for all regression models is a critical limitation of the current results. The reported unscaled RMSE, MAE, and MAPE values are unreliable and prevent a direct, meaningful comparison of forecast errors in the original inflation units across all models (including ARIMA). This unscaling issue must be resolved or clearly acknowledged as a major caveat in the final report.

- **Model Choice Recommendation (with Caveats):** Based purely on the predictive accuracy observed on the scaled test data, LASSO appears marginally better than Elastic Net. However, Elastic Net (with the optimal $\alpha = 0.9$) provides very similar performance and might be theoretically preferred for its robustness when predictors are correlated. The choice between these two is less critical than addressing the fundamental unscaling problem. Ridge Regression and the standard univariate ARIMA model are clearly less suitable for forecasting Headline CPI based on these results.

5.5.3 Recommendations for Future Work

Based on the findings and limitations of this study, the following recommendations are made for future work:

- ****Resolve Prediction Unscaling Issue:**** Prioritize investigating and correcting the procedure for unscaling predictions from the regression models to enable reliable evaluation and comparison of forecast errors in the original inflation units.
- ****Implement ARIMAX Models:**** Explore ARIMAX models, which combine the time series component of ARIMA with the ability to incorporate external regressors. This could potentially capture both the historical time dynamics and the influence of key macroeconomic predictors.
- ****Conduct Rolling Window Evaluation:**** Implement a rolling window forecasting evaluation to assess model performance more robustly across different time periods and economic conditions, rather than relying on a single train/test split.
- ****Explicitly Model Structural Breaks:**** Investigate and potentially model structural breaks in the inflation series or the relationships between variables, as these are common in volatile economies like Pakistan and can significantly impact forecasting accuracy.
- ****Explore Alternative Predictor Specifications:**** Further refine the set of independent variables and their lags based on economic theory and advanced variable selection techniques.

References

- [1] State Bank of Pakistan (sbp.org.pk). *Evaluating the Performance of Inflation Forecasting Models of Pakistan*.
- [2] DergiPark (dergipark.org.tr). *Forecasting Inflation through Econometric Models: An Empirical Study on Pakistani Data*.

- [3] IDEAS/RePEc (ideas.repec.org). *Evaluating the Performance of Inflation Forecasting Models of Pakistan.*
- [4] Ministry of Finance, Government of Pakistan (finance.gov.pk). *Inflation.*
- [5] ResearchGate (researchgate.net). (PDF) *Three Attempts at Inflation Forecasting in Pakistan.*
- [6] EconStor (econstor.eu). *Modeling and forecasting Pakistan's inflation by using time series ARIMA models.*
- [7] University of the Punjab (pu.edu.pk). *INFLATION FORECASTING IN PAKISTAN USING ARTIFICIAL NEURAL NETWORKS.*
- [8] ResearchGate (researchgate.net). *P-Star Model: A Leading Indicator of Inflation for Pakistan.*
- [9] PIDE (pide.org.pk). *Inflation in Pakistan: High-Frequency Estimation and Forecasting.*
- [10] PIDE (pide.org.pk). *Inflation Forecasting for Pakistan in a Data-rich Environment (Article).*
- [11] IMF Working Paper (elibrary.imf.org). *Three Attempts at Inflation Forecasting in Pakistan.*
- [12] Munich Personal RePEc Archive (mpra.ub.uni-muenchen.de). *Evaluating Performance of Inflation Forecasting Models of Pakistan.*
- [13] CiteSeerX (citeseerx.ist.psu.edu). *Forecasting inflation and economic growth of Pakistan by using two time series methods.*
- [14] EconPapers (econpapers.repec.org). *EconPapers.*
- [15] IDEAS/RePEc (ideas.repec.org). *Inflation Forecasting for Pakistan in a Data-rich Environment (Article).*
- [16] ResearchGate (researchgate.net). *Model Specification and Inflation Forecast Uncertainty in the Case of Pakistan.*
- [17] arfjournals.com. *www.arfjournals.com.*
- [18] ResearchGate (researchgate.net). (PDF) *Forecasting Inflation, Exchange Rate, and GDP using ANN*
- [19] Google Scholar (scholar.google.com). *Waseem Khoso.*
- [20] ResearchGate (researchgate.net). *Inflation Forecasting in Turbulent Times — Request PDF.*
- [21] Munich Personal RePEc Archive (mpra.ub.uni-muenchen.de). *Browse by Languages.*
- [22] IDEAS/RePEc (ideas.repec.org). *Forecasting Inflation Applying ARIMA Model with GARCH Innovation*

- [23] CSRC Publishing (publishing.globalcsrc.org). *Forecasting Inflation Applying ARIMA Model with GARCH Innovation: The Case of Pakistan*.
- [24] UTB (utb.edu.bh). *AMAIUB Research Catalogue*.
- [25] ResearchGate (researchgate.net). *Forecasting Inflation in Developing Nations: The Case of Pakistan*.
- [26] IDEAS/RePEc (ideas.repec.org). *Forecasting Inflation by Using the Sub-Groups of both CPI and WPI*
- [27] IBA (iba.edu.pk). *A Machine Learning Guide for Econometricians – A Hands-on*
- [28] MDPI (mdpi.com). *Forecasting of Inflation Based on Univariate and Multivariate Time Series Models: An Empirical Application*.
- [29] ResearchGate (researchgate.net). *Forecasting of Inflation Based on Univariate and Multivariate Time Series Models: An Empirical Application*.
- [30] Taylor & Francis Online (tandfonline.com). *Full article: Forecasting Inflation in Mongolia Using Machine Learning*.
- [31] Lahore School of Economics (lahoreschoolofeconomics.edu.pk). *lahoreschoolofeconomics.edu.pk*.
- [32] IDEAS/RePEc (ideas.repec.org). *Macroeconomic forecasting for Pakistan in a data-rich environment*.
- [33] ResearchGate (researchgate.net). *Macroeconomic forecasting for Pakistan in a data-rich environment*.
- [34] PIDE (file.pide.org.pk). *Inflation in Pakistan: High-Frequency Estimation and Forecasting*.
- [35] Taylor & Francis Online (tandfonline.com). *Macroeconomic forecasting for Pakistan in a data-rich environment*.
- [36] ResearchGate (researchgate.net). *Forecasting the inflation in Pakistan; the Box-Jenkins approach*
- [37] PalArch's Journal of Archaeology of Egypt/Egyptology (archives.palarch.nl). *Vol. 18 No. 08 (2021)*.
- [38] CSRC Publishing (publishing.globalcsrc.org). *Forecasting Inflation, Exchange Rate, and GDP using ANN and ARIMA Models: Evidence from Pakistan*.
- [39] University of Gujrat (uog.edu.pk). *Annual Report 2020-21*.
- [40] Federal Reserve Board (federalreserve.gov). *Forecasting US inflation in real time*.
- [41] Scholars at Harvard (scholar.harvard.edu). *Forecasting Inflation*.
- [42] IMF (imf.org). *Three Attempts at Inflation Forecasting in Pakistan – Madhavi Bokil and Axel Schimmelpfennig – IMF Working Paper 05/105, May 1*.

- [43] Munich Personal RePEc Archive (mpra.ub.uni-muenchen.de). *Evaluating Performance of Inflation Forecasting Models of Pakistan*.
- [44] Niner Commons (ninercommons.charlotte.edu). *EVALUATING THE IMPACT OF MACRO ECONOMIC VARIABLES ON INFLATION AND FORECASTING INFLATION* by Rachana Pandey.
- [45] NBER (nber.org). *NBER WORKING PAPER SERIES FORECASTING INFLATION* James H. Stock NATIONAL BUREAU OF ECONOMIC RESEARCH March 1999.