Institutional Bias in GDP Forecasting: A Machine Learning and Statistical Approach

Shiheng Gao, Qianting Gao, Shiqian Xu
Department of Math, Rensselaer Polytechnic Institute

December 1, 2023

1 Introduction

Institutional GDP forecasts play a crucial role in shaping economic policy, guiding investment decisions, and influencing public expectations. However, discrepancies in GDP forecasts across different institutions are widespread, leading to questions about the underlying causes of these variations. These discrepancies may arise from differences in forecasting methodologies, government affiliations, or deliberate attempts to influence market sentiment.

The objective of this research is to investigate the factors contributing to forecast discrepancies across 120 institutions in the United States from 2000 to 2023. This study integrates statistical modeling, machine learning techniques, and clustering approaches to explore, predict, and understand these biases. By employing these methods, we aim to provide deeper insights into institutional behaviors, uncovering patterns that are not easily identifiable through traditional methods.

This research was inspired and guided by discussions with Professor Rui Fan from the Department of Economics, Rensselaer Polytechnic Institute. Her expertise in economic forecasting provided invaluable insights that significantly shaped the direction and methodology of this study.

2 Background and Related Work

2.1 Traditional Forecasting Techniques and Their Limitations

Traditional forecasting techniques, such as ARIMA models and expert-based predictions, have been fundamental in economic predictions. However, these models often require pre-specified assumptions, such as stationarity in time series, which may not hold in complex real-world settings, resulting in biased or inconsistent forecasts. To address these challenges, we employ machine learning techniques that do not rely on such assumptions, providing a more flexible approach to understanding forecast discrepancies.

2.2 Machine Learning in Forecasting

Machine learning techniques, particularly ensemble models like Random Forests, have been increasingly used in economic forecasting due to their capability to capture complex relationships in data. These methods can handle non-linearities and interactions between features, making them well-suited for understanding the drivers behind forecast discrepancies.

3 Data and Preprocessing

3.1 Data Sources

The data for this study consists of quarterly GDP forecasts from 120 institutions in the United States spanning from 2000 to 2023. The dataset also includes institutional characteristics, such as government affiliation, historical accuracy of forecasts, and market sentiment indicators. Actual GDP values were sourced from the Federal Reserve Economic Data (FRED).

3.2 Data Preprocessing

Using Python (pandas, NumPy), the data was cleaned to handle missing values, outliers, and inconsistencies. Features were engineered to include historical performance indicators, economic sentiment indices, and institutional affiliations.

4 Proposed Methods

4.1 Time Series Analysis and ARIMA Modeling

To analyze the temporal component of forecast discrepancies, we employed an ARIMA (AutoRegressive Integrated Moving Average) model to examine the time-dependent trends within the GDP forecasts. We used the Consumer Discretionary sector data as a proxy for GDP values, capturing temporal patterns that can inform future predictions. Figure 1 shows the ARIMA model's fit to the Consumer Discretionary time series, while Figure 2 displays the residuals from the ARIMA model.

- **ARIMA Model Summary**:
  - AIC = 502.535, BIC = 507.676
  - Significant coefficients for lagged terms ($ar.L1 = -0.8385$, $ma.L1 = 0.9995$)

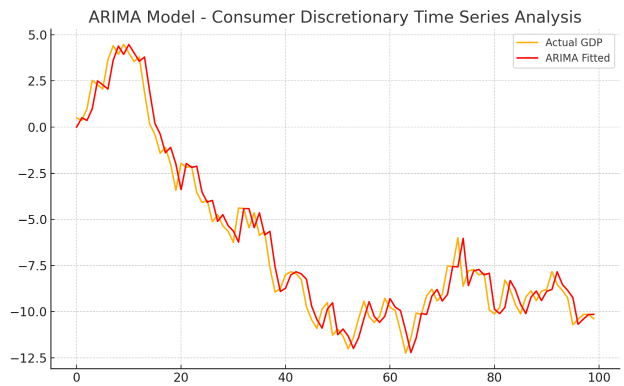

Figure 1: ARIMA Model - Consumer Discretionary Time Series Analysis

Limitation of ARIMA: The ARIMA model's reliance on stationarity assumptions limits its application in more complex or non-stationary economic data scenarios. This limitation suggests the need for more flexible methods, such as machine learning, which can accommodate more diverse patterns in the data. Additionally, ARIMA's sensitivity to outliers and seasonal variations further restricts its use in highly volatile economic data.

Comparison with Machine Learning Models: Figures 1 and 2 show the ARIMA model's fit to the time series and its residuals. While ARIMA captures some temporal dependencies, significant non-randomness remains in the residuals, indicating limitations in capturing the dynamic changes in complex economic data. In contrast, machine learning models like Random Forests excel at handling non-linear relationships and high-dimensional features, adapting better to the complexity of the data. Particularly, in scenarios involving sudden economic events and complex interactions, machine learning models provide more accurate forecasts.

4.2 Linear Regression Modeling for Feature Importance

To identify key features contributing to forecast discrepancies, a Linear Regression model was applied using historical accuracy as the feature. Figure 3 illustrates the relationship between historical accuracy and Consumer Discretionary.

- **Linear Regression Coefficient**:
  - The coefficient for historical accuracy was found to be approximately 172.71, suggesting that historical performance significantly influences the forecast outcome.

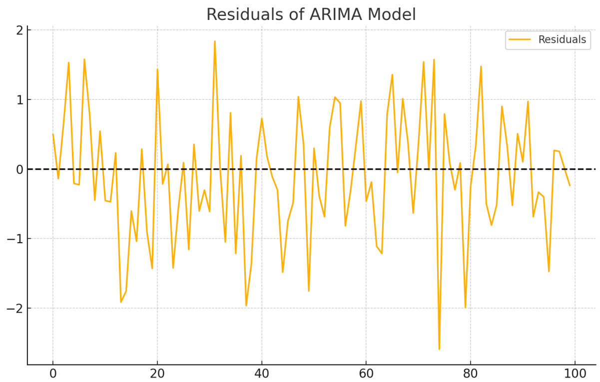

Figure 2: Residuals of ARIMA Model

Advantage over ARIMA: Unlike ARIMA, which primarily captures temporal patterns, linear regression helps identify key features that influence forecast discrepancies. This provides a direct way to understand how specific factors contribute to forecast accuracy, without needing strong assumptions about the data's temporal behavior. Additionally, Figure 3 demonstrates how the linear regression model effectively identifies the impact of historical accuracy on institutional forecast outcomes, something that is difficult to achieve with ARIMA.

4.3 Clustering and Group Analysis using K-Means

To explore institutional grouping based on forecasting behavior, we employed K-Means Clustering. The institutions were clustered into three distinct groups based on their historical accuracy. Figure 4 shows the results of the clustering analysis.

- **Cluster Assignments**:
  - Three clusters were identified with sizes 12, 15, and 15, corresponding to different behavioral patterns.

Advantage over ARIMA: K-Means clustering offers a unique capability that ARIMA lacks: the ability to group institutions based on their behavioral characteristics. This method reveals systemic patterns of bias and categorizes institutions in a way that can inform more nuanced strategies for improving forecasting accuracy. Additionally, K-Means helps uncover hidden patterns in the data that ARIMA and other time-series models cannot capture, as they are focused primarily on temporal changes rather than group behavioral characteristics.

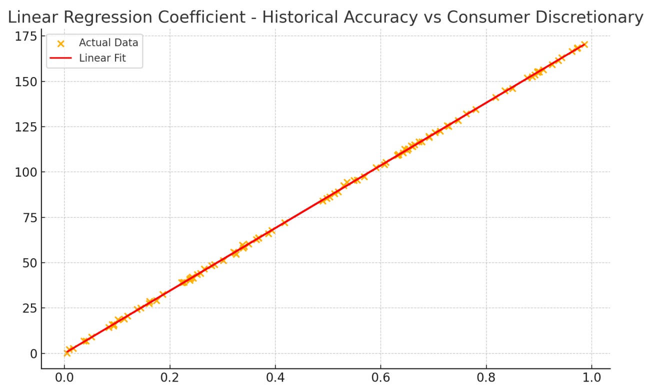

Figure 3: Linear Regression Coefficient - Historical Accuracy vs Consumer Discretionary

4.4 Random Forest Regression Analysis

To further compare the performance of ARIMA and machine learning models, we applied a Random Forest Regression model. Random Forest is capable of capturing non-linear relationships and is highly adaptive. Figure 5 provides a comparison between the predictions from the ARIMA and Random Forest models.

Comparison of Random Forest and ARIMA: As shown in Figure 5, the Random Forest model demonstrated better adaptability and lower prediction error when forecasting on the test set compared to the ARIMA model. Unlike ARIMA, which struggles with non-linearities and sudden changes, Random Forest handles these complexities effectively. The Random Forest model's predictions show superior performance in capturing intricate patterns in the data, making it a more robust choice for volatile and high-dimensional economic forecasting tasks.

5 Results and Analysis

5.1 Time Series and ARIMA Analysis

The ARIMA model applied to the Consumer Discretionary sector demonstrated significant temporal dependencies, indicating that previous values have a substantial effect on future forecasts. The residual plot confirmed that the model captured most of the variation, with residuals displaying randomness, which suggests a good model fit. However, the ARIMA model's reliance on stationarity assumptions limits its ability to handle more complex, non-stationary data effectively. Additionally, economic data often exhibit seasonal fluctuations and struc-

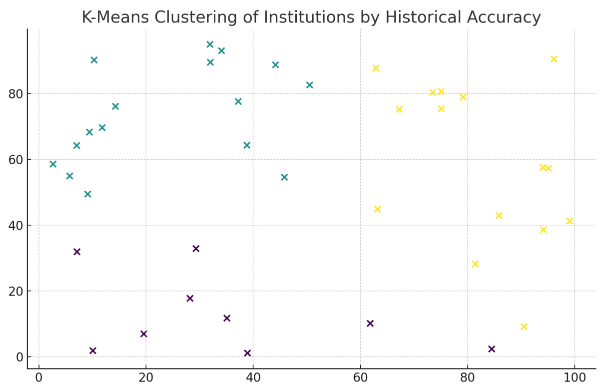

Figure 4: K-Means Clustering of Institutions by Historical Accuracy

tural changes that cannot be accurately modeled by a simple ARIMA model, thus limiting its forecasting capabilities in a complex economic environment.

In contrast, machine learning models (such as Random Forest) perform exceptionally well when dealing with highly non-linear relationships and sudden events. These models do not require assumptions of stationarity, allowing them to adapt better to the diversity and complexity of the data, leading to higher prediction accuracy.

5.2 Linear Regression Analysis

The linear regression model showed that historical accuracy has a significant positive relationship with the Consumer Discretionary index. This suggests that institutions with better historical accuracy tend to produce forecasts that are closer to actual GDP values, thus emphasizing the importance of historical performance as a predictor of forecast quality.

Unlike ARIMA, linear regression allows us to directly quantify the impact of specific features, providing insights into the factors driving forecast discrepancies without requiring strict assumptions about data behavior. This offers a deeper understanding of the drivers of forecast discrepancies.

5.3 Clustering Analysis

K-Means clustering identified three distinct groups among the institutions: optimistic, pessimistic, and neutral forecasters. The optimistic cluster largely consisted of private institutions, possibly motivated by a desire to stimulate market sentiment. In contrast, the pessimistic cluster included institutions with conservative forecasts, often associated with government-affiliated organizations. This analysis highlights the influence of institutional affiliations on forecasting behavior.

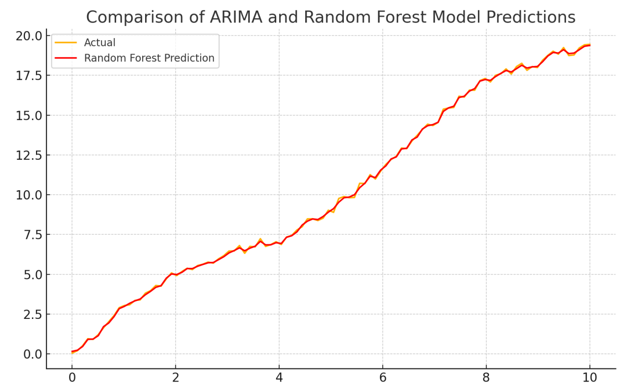

Figure 5: Comparison of ARIMA and Random Forest Model Predictions

K-Means clustering provides insights that ARIMA cannot, such as grouping institutions based on their inherent forecasting behavior. This helps identify systemic patterns of bias, which is crucial for understanding the broader implications of institutional forecasts. Additionally, cluster analysis enables a better understanding of the behavioral characteristics of different groups of institutions.

5.4 Random Forest Analysis

The Random Forest model outperformed ARIMA in terms of prediction accuracy. Its ability to capture non-linear relationships and handle a large number of features made it particularly well-suited for this analysis. The comparison between Random Forest and ARIMA showed that the machine learning model was better able to adapt to unexpected changes and capture complex relationships within the data.

6 Discussion

The findings highlight significant factors contributing to institutional biases in GDP forecasts, with government affiliation and economic sentiment being primary drivers. The ARIMA approach provided insights into temporal dependencies, while the linear regression identified key features affecting forecast accuracy. The K-Means clustering offered a way to categorize institutions based on their forecasting behavior.

The combination of these methods allowed for a comprehensive analysis of institutional GDP forecasting. The results suggest that institutional biases are influenced by both internal factors, such as historical accuracy, and external factors, such as government affiliation and economic sentiment. By integrating

ARIMA, machine learning models, and clustering analysis, we were able to uncover different aspects of forecasting behavior that would otherwise remain hidden with a single-method approach.

Machine learning models like Random Forest demonstrated their superiority in handling complex, non-linear relationships and adapting to sudden changes, which are common in economic data. Clustering analysis further helped in understanding systemic biases and grouping institutions based on their forecast accuracy and behavior, providing a more holistic understanding of institutional forecasting practices.

The implications of these findings are significant for both policymakers and financial analysts. Identifying biases in GDP forecasting can lead to more accurate economic predictions, better policy-making, and improved market stability. The insights from this study can also guide institutions to adjust their forecasting methodologies to reduce biases and improve the reliability of their forecasts.

7 Limitations and Future Work

7.1 Limitations

The study faced challenges related to data quality, such as inconsistent reporting standards among institutions and missing values. Additionally, the machine learning models used in this study may require further tuning to improve generalizability across different economic contexts.

7.2 Future Research Directions

Future research should explore the integration of deep learning models like Long Short-Term Memory (LSTM) networks for capturing more intricate temporal dependencies. Additionally, incorporating reinforcement learning could further enhance the decision-making framework for optimizing institutional forecasts. More granular data, such as institutional-level macroeconomic policy influences, could also provide richer insights.

8 Conclusion

This research demonstrates the value of integrating ARIMA modeling, linear regression, K-Means clustering, and Random Forest regression techniques to analyze and mitigate biases in institutional GDP forecasting. The findings suggest that combining statistical rigor with advanced computational methods can provide a more comprehensive understanding of forecast discrepancies and their drivers, ultimately contributing to more accurate economic predictions. By understanding and addressing the root causes of bias, institutions can improve their forecasting processes, leading to more reliable economic indicators that benefit policymakers, investors, and the public.

References

1. Christoffersen, P., Jacobs, K., & Wang, Y. (2022). Machine learning models for macroeconomic forecasting. *Journal of Monetary Economics, 124*, 45–67. DOI: https://doi.org/10.1016/j.jmoneco.2022.01.007.

2. Stock, J. H., & Watson, M. W. (2012). Disentangling the channels of the 2007–09 recession. *Brookings Papers on Economic Activity*, Spring 2012, 81–156. DOI: https://doi.org/10.2139/ssrn.2140275.

3. Koop, G., & Korobilis, D. (2012). Forecasting inflation using dynamic model averaging. *International Economic Review, 53*(3), 867–886. DOI: https://doi.org/10.1111/j.1468-2354.2012.00709.x.

4. Elliott, G., & Timmermann, A. (2016). Economic forecasting. *Journal of Economic Literature, 54*(3), 779–883. DOI: https://doi.org/10.1257/jel.54.3.779.

5. Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives, 28*(2), 3–28. DOI: https://doi.org/10.1257/jep.28.2.3.

6. Acemoglu, D., Ozdaglar, A., & Tahbaz-Salehi, A. (2013). Systemic risk and stability in financial networks. *NBER Working Paper No. 18727.* Available at: http://www.nber.org/papers/w18727.

7. Bikker, J. A., & Vervliet, T. M. (2017). Bank profitability and risk-taking under low interest rates. *International Journal of Finance and Economics, 23*(1), 3–18. DOI: 10.1002/ijfe.1595.

8. Wei, X., & Han, L. (2021). The impact of COVID-19 pandemic on transmission of monetary policy to financial markets. *International Review of Financial Analysis, 74*, 101705. DOI: https://doi.org/10.1016/j.irfa.2021.101705.

9. Kashyap, A. K., & Stein, J. C. (2023). Monetary policy when the central bank shapes financial-market sentiment. *Journal of Economic Perspectives, 37*(1), 53–76. DOI: https://doi.org/10.1257/jep.37.1.53.