Moving With the Times: Updating the Central Bank's Nowcasting Framework for Relevancy in a Post-COVID-19 World

Marco A. Chan

Research and Economic Analysis Department

Central Bank of Belize

The views expressed in this research paper are those of the author and do not represent the views of the Central Bank.

Correspondence concerning this article should be addressed to Marco Chan, Research and Economic Analysis Department, Central Bank of Belize, Gabourel Lane, Belize City, Belize.
Email: Marco.Chan@centralbank.org.bz.

Abstract

This paper improves the quarterly Gross Domestic Product (GDP) nowcasting framework used by the Central Bank of Belize to account for structural changes in Belize's economy brought on by a rebasing exercise and the volatility induced by the COVID-19 pandemic. Employing a bottom-up approach, the framework utilises 16 Autoregressive Integrated Moving Average with exogenous variables (ARIMAX) models, employed using the Box-Jenkins (1968) approach, to nowcast industrial value-added components. These are then aggregated to derive the quarterly GDP estimate. The exogenous variables include timely industrial indicators and dummy variables, strategically incorporated to capture the economic fluctuations caused by the pandemic. The results demonstrate a significant improvement in the Mean Absolute Percentage Error (MAPE) for both the pre- and post-pandemic periods. Specifically, the MAPE decreases from 4.1% to 1.7% for 2017Q1 – 2019Q4 and from 8.5% to 2.4% 2020Q1 – 2022Q2, highlighting the enhanced accuracy of the new framework compared to benchmark models. The findings indicate that ARIMAX models offer superior out-of-sample forecasting accuracy and provide valuable insights for policymakers navigating economic shocks and structural adjustments.

Keywords: nowcasting, pandemic, ARIMAX, Box-Jenkins, industrial models, dummy variables, MAPE, RMSE

Table of Contents

1.0 Introduction

Since 2017, the Central Bank of Belize has been using a nowcasting framework consisting of 11 linear dynamic industrial models and 34 bridge indicators to estimate Belize's GDP quarterly. Initially developed by Arana (2015), this framework allowed the Bank to produce quarterly GDP estimates within five weeks of a quarter's end, providing stakeholders with a timelier view of economic conditions. In comparison, the Statistical Institute of Belize (SIB) produces its first round of quarterly GDP estimates nearly three months after the quarter under review.

The need to update the nowcasting framework was motivated by the impact of SIB's GDP rebasing exercise in 2022. The rebasing exercise, which used the System of National Accounts (SNA) 2008 compared to SNA 1993, identified seven new industries, rendering the old nowcasting framework obsolete (SIB, 2020). Additionally, the nowcasting model's accuracy plummeted in the wake of the COVID-19 crisis, with the MAPE more than doubling from 4.1% to 8.7%. Its sectoral components, anchored on ordinary least squares (OLS) and autoregressive integrated moving average (ARIMA) models, could not predict the erratic shifts in economic conditions. Consequently, the model's efficacy in forecasting short-term GDP diminished, leading to higher margins of error and reduced reliability of the overall GDP estimate.

To address these challenges, this paper aims to enhance the Central Bank's GDP nowcasting framework to better capture the structural changes, arising from the GDP rebasing exercise, and to model the sharp economic fluctuations, stemming from the COVID-19 shock. This involved an overhauling of the existing 11 sectoral model framework to utilise

16 Autoregressive Integrated Moving Average with exogenous variables (ARIMAX) models, employing the Box-Jenkins (1968) approach. To address the heterogeneous response to the pandemic across industrial categories, dummy variables are strategically incorporated to improve model accuracy. The exogenous variables were selected in accordance with the SIB's "*System of National Accounts Documentation*" (2020).

A comprehensive review of the previous and proposed frameworks' performance revealed improvements in the accuracy of quarterly GDP estimates, with the MAPE improving from 6.1% between 2017Q1 and 2022Q1, to 2.1% between the same time period. The enhanced precision of the new nowcasting framework should provide the Central Bank with timely insights into the economic landscape and reveal nuanced changes across sectors over time. This granular understanding will enable the identification of sector-specific trends and serve as an early warning system for potential economic imbalances.

The paper is structured as follows: Section II reviews the pertinent literature. Section III covers the methodology employed in constructing the ARIMAX models. Section IV presents empirical results, while Sections V and VI provides a discussion on the results generated and conclusions, respectively.

2.0 Literature Review

2.1 Nowcasting and Evolution of Techniques

Nowcasting, or the prediction of current economic conditions based on available data, plays a crucial role in informing timely policy decisions and providing stakeholders with up-to-date assessments of economic performance. The importance of accurate and efficient nowcasting frameworks has been underscored by recent events such as the COVID-19 pandemic, which caused unprecedented volatility and structural shifts in economies worldwide. In recent years, there has been a growing body of research focused on developing and refining nowcasting methods. Studies have explored various approaches, including bridge models, dynamic factor models, and mixed-frequency data analysis, to enhance the accuracy and timeliness of GDP estimates. Since Giannone, Reichlin, & Small (2006) introduced a formal process for updating nowcasting frameworks, numerous techniques and tools have been developed to improve the accuracy and timeliness of estimates.

2.2 Bottom-Up Approach to Nowcasting

Previously, the Central Bank relied solely upon a bottom-up approach, aggregating 11 sectoral bridge models to produce a quarterly GDP estimate based on the work of Arana (2015). Although Arana (2015) preferred indicator models, sectoral models were used due to limits on data availability and the added benefit of observing nuanced trends within GDP's various components through the disaggregation, as was found to be the case in Kaustubh, Bhadury, & Ghosh (2024) and Dias, Pinhiero, & Rua (2016).

Several studies suggest that a bottom-up approach to nowcasting GDP components for subsequent aggregation can enhance the accuracy of estimates. For instance, Dias, Pinheiro, and Rua (2016) employed factor models to project expenditure-side components of GDP and found that performance outstripped that of univariate benchmark models. Similarly, Hahn and Skudelny (2008) adopted a production-side approach to forecast GDP sub-categories, demonstrating not only improved accuracy but also the valuable insights gained from understanding the co-movements of underlying drivers of economic growth. Moreover, the bottom-up approach also provided a means for accounting for sector-specific shocks since abnormal growth levels in one component of GDP could be easily identifiable and modelled (Hahn & Skudelny, 2008). This characteristic was explored following the pandemic as researchers sought to capture the impact of the pandemic across various economic sectors. Kaustubh, Bhadury, and Ghosh (2024) produced a production-side nowcast in which the impact of the pandemic was accounted for within each sub-categorical estimate. This was done using dummy variables representing the onset of the pandemic, which were found to be statistically insignificant for select subcategories and highly significant for others (Kaustubh, Bhadury, & Ghosh, 2024).

2.3 Challenges to Nowcasting: Structural Changes and Economic Shocks

Relationships between indicators and GDP can change over time, owing mainly to large structural shifts, as noted in Hahn and Skudelny (2008). In developing countries, rebasing exercises, often result in stark changes to the fundamental structure of an economy and underlying drivers of growth. For example, Akpan & Udofia (2017) found that Nigeria's largest contributor to growth had shifted away from manufacturing to services in their study following a rebasing exercise. The consequences of this were profound for Nigeria, as was also the case for Belize.

Bridge and mixed data sampling (MIDAS) models, in particular, failed to capture the magnitude of impact on GDP due to their primary premise of capturing the average behaviour of variables over a period of time (Foroni, Marcellino, & Stevanovic, 2022). This was widespread as Arana (2015), Cascaldi-Garcia, Luciani, & Modugno (2023), Kaustubh, Bhadury, & Ghosh (2024), and others pointed out the ephemeral nature of indicator explanatory power. The onset of the COVID-19 pandemic also severely disrupted the effectiveness of sectoral models and other traditional forecasting methods, as they failed to capture the impact of sharp fluctuations on GDP (Cascaldi-Garcia, Luciani, & Modugno, 2023).

2.4 Innovations in Nowcasting in Response to the Pandemic

Incidentally, the pandemic catalysed growth in research within the field of nowcasting resulting in a growing body of work on the subject matter. The focus of these studies were multifaceted, ranging from the recalibration of models and frameworks using more sophisticated econometric tools as was the case in Cascaldi-Garcia, Luciani, & Modugno (2023), to the reworking of simpler approaches such as Dias, Pinhiero, & Rua (2016), as well as the incorporation of alternative data sources like Nakazawa (2022) and Kaustubh, Bhadury, & Ghosh (2024). The pandemic severely impacted the ability of nowcasting models to assess the severity of the economic downturn. This was stated by Foroni, Marcellino, & Stevanovic (2022) and Nakazawa (2022), in observance of their respective bridge and unrestricted mixed data sampling (UMIDAS) models, which failed to predict the depth of the downturn due to a lack of timely, statistically significant data. The same holds true for the more sophisticated Dynamic Factor Models (DFM), which were widely employed during the pandemic.

Most models were not equipped to capture the magnitude of the pandemic without significant adjustment through the incorporation of alternative high-frequency data sources such as in Cascaldi-Garcia, Luciani, & Modugno (2023) and Kaustubh, Bhadury, & Ghosh (2024). Machine learning techniques have demonstrated some superiority in predicting both the occurrence and depth of contraction brought on by anomalous occurrences. In Maccarrone, Morelli, & Spadaccini (2021), traditional ARIMA and SARIMAX models were pitted against a machine learning model, which emerged as the most accurate model when nowcasting one quarter ahead, with significantly smaller margins of error. Similarly, a study across six European countries, Dauphin, et al. (2022) found that machine learning models surpassed benchmark and DFM models in reducing forecast errors during the pandemic in most countries. Barrios et al. (2021) furthered the discourse in Belize by introducing six machine learning models, including lasso regressions, random forest algorithms, and neural networks. These models generate a singular quarterly GDP estimate, compiled into an ensemble figure based on weights derived from the inverse of RMSE produced by each model (Barrios, Martin, Escobar, Pena, & Leslie, 2021).

Nevertheless, the performance of traditional models cannot be discounted, as Maccarrone, Morelli, & Spadaccini (2021)'s SARIMAX models outperformed the machine learning models when looking at longer horizons and with the introduction of covariates. It is important to note that the current framework utilised by the Bank is a bottom-up approach employing 11 bridge sectoral models to produce one estimate. This approach, distinct from both Maccarrone, Morelli, & Spadaccini (2021) and Barrios et al. (2021), provides a more nuanced observance of the economy. Such approaches to nowcasting are rare but not unprecedented.

2.5 Theoretical Framework for the Current Study

The use of dummy variables to capture the impact of the pandemic on economic activity has been explored to varying degrees in the years since the outbreak. In their study of addressing outliers within their vector autoregressive (VAR) models, Carriero et. al (2024) opted to use dummy variables to capture the monthly impact of the pandemic, which absorbed the VAR residuals, improving the fitness of the model. Meanwhile, Furceri et. al (2021), utilise dummy variables within their models to capture multiple pandemics within a 20-year span and assess the most recent pandemic's impact on inequality. Nevertheless, the most relevant study that mirrors the efforts of this paper remains Kaustubh, Bhadury, and Ghosh's (2024). Their use of dummy variables to capture the pandemic's effect within a bottom-up approach to nowcasting GDP, combined with Maccarrone et al. (2021)'s ARIMAX-based approach to nowcasting forms the theoretical backbone of this paper.

3.0 Methodology

3.1 Overview

This paper aims to enhance the Bank's GDP nowcasting framework to accurately capture recent structural changes brought on by a recent rebasing exercise and effectively model the impact of the pandemic. Within this section, the application of the Box-Jenkins (1968) three-step approach allowed for the creation of 16 Auto Regressive Integrated Moving Average with Exogenous variable (ARIMAX) industrial models. These were then aggregated to nowcast quarterly constant GDP.

To capture the heterogenous effects of the pandemic on the various industries, two dummy variables were created. These variables represent the initial downturn and subsequent rebound in GDP activity during and after the pandemic[1]. The inclusion of these dummies, either individually or in combination, was informed by Chow breakpoint tests conducted for each industry. The exogenous component in each of the 16 models were then identified from the SIB's "*System of National Accounts*" documentation and assigned. Following these steps, the industrial models were estimated, with relevant diagnostics conducted.

3.2 Industrial Models

3.2.1 Selection of Industrial Groupings

Following the SIB's 2022 GDP rebasing exercise, the number of industrial classifications rose from 11 to 21, see Table A1. In this paper, the 11 additional industries were aggregated into six industrial groupings[2]. This study diverges from the previous framework's structure of modelling each industrial classification for two reasons. The first was the unavailability of timely variables. The second was lack of utility gained from nowcasting industries with miniscule contributions to GDP. The Table below identifies the new industrial groupings.

---

[1] Q1 2020 – Q4 2021

[2] Due to the absorption of subcategory "Fishing" into "Agriculture, Forestry and Fishing", only 10 of the initial 11 industrial categories remained, with 11 new categories identified.

Table 1: Components of Industrial Grouping

Industrial Components	Grouping Name
Mining	Mining
Water Supply	Water Supply
Real Estate Activities	Real Estate Activities
Professional Scientific and Technical Activities	Professional and Adminstrative Activities
Adminstrative and Support Service Activities	
Education	Education and Health Activities
Human Health and Social Work Activities	
Information and communication	Other Activities
Arts,entertainment, recreation	
Other service activities	
Activities of households as employers	

The grouping of eight industries together to form three categories resulted from an investigation into the explanatory power of the selected exogenous variables in determining the gross value added (GVA) estimates of the industries. Given that variables for the six industries were not available in a timely manner, the suitability of various alternate quarterly and monthly indicators were assessed via correlation testing for the eight industries. This was done according to the SIB's guidelines provided in the "*System of National Accounts Documentation*" (2020), which outlines the methods and indicators used to compile GDP and its industries. From here, industries that shared explanatory variables were combined.

As a result, "*Professional Scientific and Technical Activities*" and "*Administrative and Support Services Activities*" were aggregated into "*Professional and Administrative Services*", with GST, business tax, and BPO inflows proven to be significantly correlated variables. Similarly, "*Education*" and "*Human Health and Social Work Activities*" were combined into "*Education and Health Activities,*" sharing current fiscal expenditure as an explanatory variable. The largest amalgamation was of the four categories "*Information and Communication,*" "*Arts, Entertainment, and Recreation,*" "*Other Service Activities,*" and "*Activities of Households as Employers*" into "*Other Services*", which found tourism arrivals to be the most appropriate variable.

3.2.2 Data Selection

For the 16 proposed industrial models, 42 exogenous explanatory variables were used. All indicators were used by the SIB, as documented in their "*System of National Accounts*" (SIB, 2020) documentation, to generate estimates of value-added output across the industries. As shown in Table A2, 12 variables were used to directly nowcast 10 industrial models, while 30 were inputs for five indices used to nowcast the remaining six models. The indices were constructed in accordance with specifications provided by the SIB and are standardised to the base year of 2014. The construction of these indices arose in an attempt to avoid the issue of overfitting the models.

Table 2: Industrial Models Utilising Indices as Exogenous Variables

Index	**Industrial Model**
Agricultural Index	Agriculture, Forestry and Fishing
Manufacturing Index	Manufacturing
Tourist Arrival Index	Hotels and Restaurant Activities
	Other Service Activities
Transportation Index	Transportation
Professional and Administrative Services Index	Professional and Adminstrative Activities

3.2.3 Dummy Variables and Chow Breakpoint Test

As previously mentioned, to capture the effects of the pandemic on the 16 industrial categories two dummy variables were created. The first, "COVID 2020" captures the initial downturn in economic activity between the first quarter of 2020 and the first quarter of 2021. The second, "COVID 2021" encapsulates the period of high GDP growth noted between the second quarter of 2021 and first quarter of 2022. To determine the placement of the variables, a Chow Breakpoint Test was conducted on each of the 16 industrial groupings to determine when each industry specific breakpoint occurred. The equation is formed as:

$$\frac{RSS-(RSS_1+RSS_2)}{RSS_1+RSS_2}\ X\ \frac{T-2k}{k} \tag{1}$$

Where RSS is the sum of squared residuals of the sample series, $RSS_1$is the sum of squared residuals before and up to the identified break date, RSS_2 is the sum of squared residuals at and after the identified break date, T is the number of observations, and k is the number of regressors in the equation. The results of this test, dictating which dummy variables were included in each industrial model is found below in Table 3.

Table 3: Results of Chow Breakpoint Test

Industrial Groupings	**COVID2020**	**COVID2021**
Agriculture, Forestry and Fishing	☐	☐
Mining	☑	☐
Manufacturing	☐	☐
Electricity	☑	☐
Water Supply	☑	☑
Construction	☑	☑
Wholesale and Retail Trade	☑	☑
Hotels and Restaurant Activities	☑	☐
Transportation	☐	☑
Financial Activities	☑	☑
Real Estate Activities	☑	☑
Professional and Adminstrative Activities	☑	☑
Public Administration and Defence Activities	☑	☑
Education and Health Activities	☐	☐
Other Service Activities	☐	☑
Taxes and Subsidies	☑	☑

3.3 ARIMAX Model and Components

An extension of the ARIMA model, the ARIMAX model incorporates the influence of one or more exogenous variables in estimating forecasts. The three parameters of the ARIMA model (p, d, q) are as follows:

The autoregressive $AR(p)$ component is defined as:

$$Y_t = \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \cdots + \emptyset_p y_{t-p} + \varepsilon_t \tag{2}$$

$$= \sum_{j=1}^{\emptyset} \emptyset_j \, y_{t-j} + \varepsilon_t$$

Where Y_t is the current observation, y_{t-j} are the past observations, $\emptyset$ is the autoregressive order of the process with coefficients $\emptyset_j$, and ε_t is the white noise component. When incorporating a lag operator, the $AR(p)$ takes the basic form:

$$y_t = \sum_{j=1}^{\emptyset} \emptyset_j L^j y_{t-i} + \varepsilon_t \tag{3}$$

Where $\emptyset(L) = 1 - \sum_{j=1}^{\emptyset} \emptyset_j L^j$ is defined as the lag polynomial that characterises the AR process. The addition of a mean to the model produces the AR model below.

$$\emptyset(L)(Y_t - \mu_t) = \epsilon_t \tag{4}$$

The moving average $MA(q)$ component of the ARIMAX model is comprises of:

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + p\theta_2 \varepsilon_{t-2} + \cdots + \theta_i \varepsilon_{t-i} \tag{5}$$

$$= \varepsilon_t + \sum_{l=1}^{q} \theta_j \, \varepsilon_{t-j}$$

$$= \theta(L)\varepsilon_t$$

Where ε_t is the current disturbance, ε_{t-j} are the past disturbances, $\theta_i$are the parameters of the MA model, and $\theta(L)$ are the moving average polynomial that characterises the MA process.

When both AR and MA processes are combined, they form the ARMA equation:

$$Y_t = \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \cdots + \emptyset_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \tag{6}$$

$$p(L)(Y_t - \mu_t) = \theta(L)\varepsilon_t$$

Thus, when incorporating the integrated order $I(d)$, an ARIMA (p, d, q) is formed as:

$$p(L)(1 - L)^d (Y_t - \mu_t) = \theta(L)\varepsilon_t \tag{7}$$

Lastly, the addition of exogenous variables reveals the final ARIMAX model form as:

$$p(L)(1 - L)^d (Y_t - \mu_t) = \beta_n x_{n,t} + \theta(L)\varepsilon_t \tag{8}$$

Where $\beta_n x_{n,t}$ is the coefficient of the exogenous variable(s).

3.4 ARIMAX Model Fitting

For the 16 ARIMAX models, the Box-Jenkins three-stage approach was implemented, namely model identification, estimation, and diagnostic checking. The 16 industrial series were found to be non-stationary, and their appropriate AR and MA orders were identified. Following this, the most appropriate model is selected for each industrial grouping after comparison utilising the Akaike

Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan-Quinn Information Criterion (HQIC). Thereafter, following forecasting of the models, the models are checked for invertibility. Lastly, to evaluate the results, the measures of forecast accuracy are set forth.

3.4.1 Identification & Estimation

The properties of the 16 industrial GVAs were analysed for stationarity through graphs and correlograms generated for each. None of the variables were found to be stationary. Following this, Augmented Dickey Fuller tests (ADF) were conducted for each to identify the level at which the variables are stationarity. The results are summarised in Table 4 below. the Autocorrelation function (ACF) and Partial Autocorrelation function (PACF) are then utilised to determine the number of MA and AR terms, respectively. This is also highlighted in Table 4.

Table 4: Differencing and AR/MA Order Results for Industrial Groups

Industrial Groupings	Differencing	AR Order	MA Order
Agriculture, Forestry and Fishing	1st	2	6
Mining	2nd	1	1
Manufacturing	1st	2	2
Electricity	2nd	2	1
Water Supply	2nd	2	1
Construction	2nd	1	1
Wholesale and Retail Trade	2nd	1	1
Hotels and Restaurant Activities	2nd	2	1
Transportation	1st	3	3
Financial Activities	2nd	2	1
Real Estate Activities	1st	5	4
Professional and Adminstrative Activities	1st	3	1
Public Administration and Defence Activities	2nd	4	3
Education and Health Activities	1st	1	1
Other Service Activities	2nd	1	1
Taxes and Subsidies	2nd	1	1

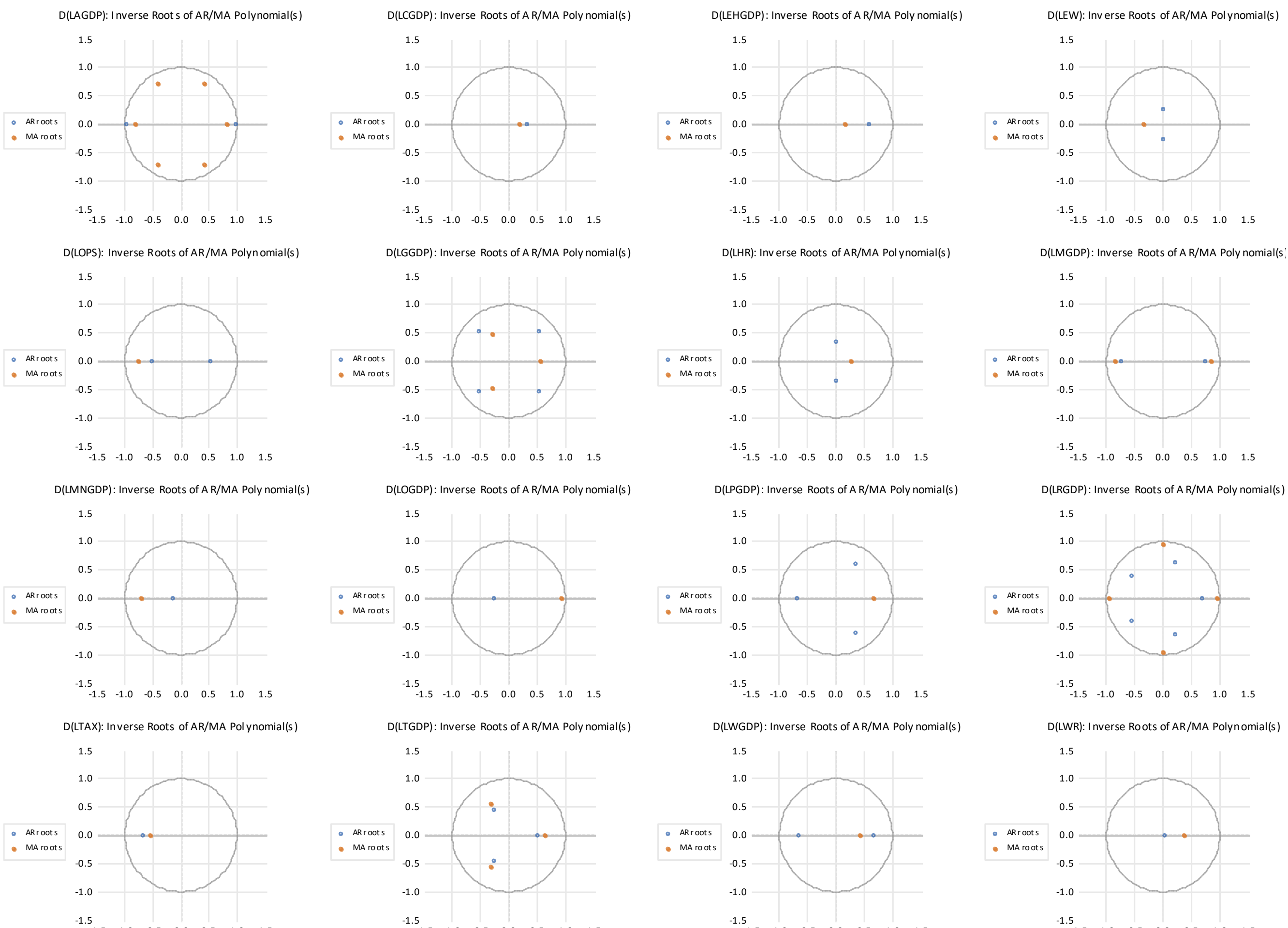
3.4.2 Estimation

Following the identification of possible suitable models for each of 16 proposed, the models were estimated using the maximum likelihood method and compared. Those with the lowest AIC, BIC and HQIC values were selected for forecasting. The AR and MA components were found to be significant across the 16 models.

3.4.3 Diagnostics

To ensure the selected models satisfy the requirements for a stable univariate process, the ACF and PACF plot of residuals are evaluated, and unit root tests are conducted. The Ljung-Box Q statistics are analysed for each of the 16 to ensure residuals are white noise. Thereafter, the results of the unit root test for serial correlation across the 16 models showed that no root of the estimated models fell out of the $-1\ and + 1$ bounds, therefore satisfying the requirement of invertibility[3]. The results of this test can be found in graph 1 below.

Figure 1: Roots of Characteristics Polynomial for 16 Industrial Models



[3] Invertibility is a desirable property for MA models in time series analysis. It ensures uniqueness, aides in estimation, an can improve forecasting performance.

3.4.4 Assessment of Forecast Accuracy

To assess forecast accuracy, three methods are utilised to test different aspects of the results of the models. The first method is the Root Mean Square Error (RMSE) statistic, which measures the difference between the predicted and estimated values by analysing the standard deviation of the residuals. It is represented by the equation:

$$\sqrt{\sum_{t=T+1}^{T+h} (\psi_t - \gamma_t)^2 / h} \tag{9}$$

Where the forecasted value is denoted by ψ_t, the actual value is γ_t, the forecast sample is $T + h$, and the period is defined as t.

The second method is the Mean Absolute Percentage Error (MAPE) approach, which represents the average of the absolute percentage errors of each entry to calculate the accuracy of forecasted values in comparison to actual values. It is represented by the equation:

$$100 \ \sum_{t=T+1}^{T+h} \left| \frac{\psi_t - \gamma_t}{\gamma t} \right| / h \tag{10}$$

Lastly, the Theil U_2 Inequality Coefficient is used to determine the forecast quality and adequacy of the models. It is represented as:

$$U_2 = \frac{\sqrt{\frac{1}{n} \sum_{t=1}^{n} (\psi_t - \gamma_t)^2}}{\sqrt{\frac{1}{n} \sum_{t=1}^{n} \gamma_t^2}} \tag{11}$$

In the first two methods, a value closer to 0 is desirable to show forecast accuracy. In the U_2 coefficient, 0 is also desirable to show accurate forecasts. However, a U_2 value equal to or above one denotes a weak model.

4.0 Results

The gross value added (GVA) of 16 industrial ARIMAX models for the period 2017Q1-2022Q4 were forecasted and aggregated to arrive at an estimate of quarterly GDP which were then compared to actual estimates. Thereafter, forecast accuracy was assessed utilising the RSME, MAPE, and Thiel U2 statistics. The utilisation of all three methods was implemented with the goal of capturing the most holistic view of the forecast performance of the models. The RMSE would indicate whether large errors are present within the forecast, while the MAPE adds the dimension of overall accuracy relative to the scale of the data. The Thiel U2 coefficient would contribute an assessment of the model's ability to capture the dynamics of change in the variable.

It was found when assessing the RMSE and MAPE that four and three of the 16 forecasts had higher than average test statistics, respectively, while a total of 7 did not fall within acceptable bounds when gauging the Thiel U2 coefficient. Interestingly, only one of these seven models were found to have experienced all three of the unfavourable statistics concurrently.

4.1 Aggregated ARIMAX Model Results

The natural logarithm values generated by the 16 industrial models were expanded and aggregated for comparison to the actual estimates. These results are summarised in the figures below. When aggregated, the RMSE for the forecast period of 2017 Q1 – 2022 Q4 was 2.6%, while the MAPE was 2.1%, indicating accurate forecasts.

Figure 2: Forecasted GDP Percent Change Versus Actual GDP Percent Change

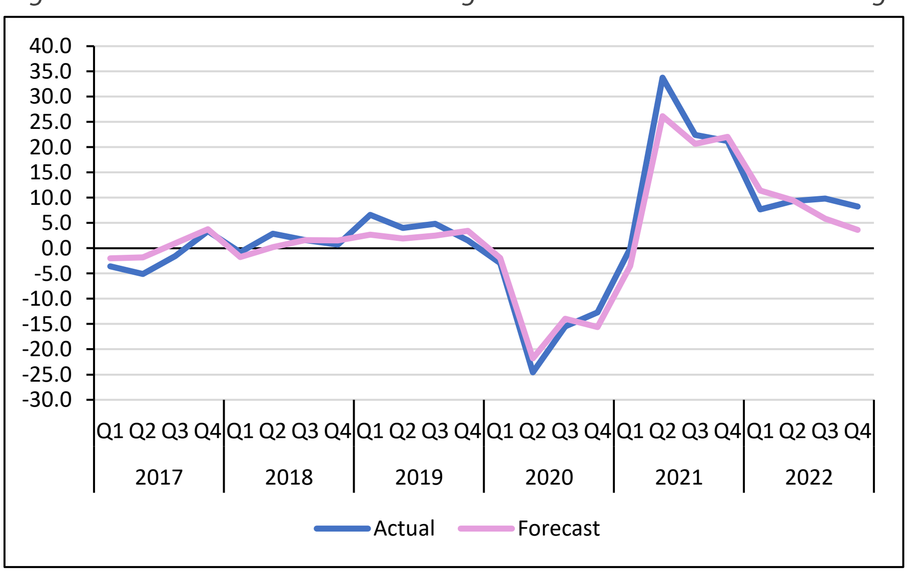
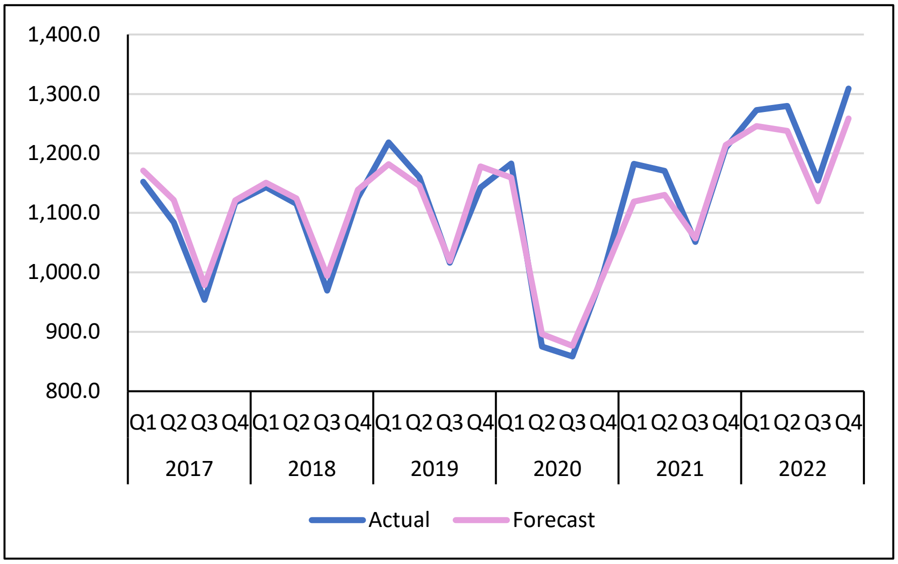


Figure 3: Forecasted GDP Value Versus Actual GDP Value



The models performed best in estimating the GVAs of the pre-pandemic era, as shown in Table 5, with the RMSE totalling 2.0% of GDP between 2017 and 2019, compared to 3.0% between 2020 and 2022. The MAPE reflected a similar narrative, averaging 1.7% in the pre-pandemic era, and 2.4% in the post-pandemic era. Volatility towards the end of the time series can be inferred from the disaggregated examination of the industrial models.

Table 5: Aggregated Annual RMSE and MAPE

	RMSE	MAPE
2017	24.5	2.0
2018	15.4	1.3
2019	26.5	1.9
2020	18.1	1.7
2021	37.9	2.5
2022	39.4	3.1

4.2 Assessment of ARIMAX Model Forecast Accuracy

The 16 ARIMAX models forecasted the period 2017Q1 to 2022Q4 and yielded results that were tested for forecast accuracy, summarised in Table 6 below. The results indicated mixed results when analysing each of three unique measures of forecast accuracy. A description of these various combinations of results is provided in Table 7.

Table 6: Measure of Forecast Accuracy by Industrial Model

Industrial Groupings	RMSE	MAPE	Thiel U2
Agriculture, Forestry and Fishing	0.18	3.14	0.35
Mining	0.24	6.95	1.04
Manufacturing	0.06	1.25	0.16
Electricity	0.20	5.88	0.68
Water Supply	0.06	2.02	1.07
Construction	0.11	2.20	1.11
Wholesale and Retail Trade	0.03	0.48	0.31
Hotels and Restaurant Activities	0.36	29.45	0.20
Transportation	0.12	2.73	0.39
Financial Activities	0.09	1.77	3.47
Real Estate Activities	0.01	0.28	2.66
Professional and Adminstrative Activities	0.16	3.33	0.77
Public Administration and Defence Activit	0.04	0.72	0.60
Education and Health Activities	0.03	0.47	0.04
Other Service Activities	0.15	2.90	1.89
Taxes and Subsidies	0.11	2.10	1.16

Table 7: Forecast Accuracy Conditions

RMSE/MAPE Condition	Thiel U2 Coefficient Condition	Result
Low RMSE/MAPE	Low Thiel U2 Coefficient	Model robustly estimates the general level and direction of a variable
Low RMSE/MAPE	High Thiel U2 Coefficient	Model estimates the general level of a variable but not the direction and magnitude of changes
High RMSE/MAPE	Low Thiel U2 Coefficient	Model estimates the direction and magnitude of changes but not the general level of the variable
High RMSE/MAPE	High Thiel U2 Coefficient	Model does not estimate the general level or the direction and magnitude well

Analysis of the RMSE, estimated from the difference between the actual and forecasted values, revealed that all models held relatively low values. While the RMSE does not have a standard range to measure results by, the general rule dictates that the closer the value is to 0, the more accurate the forecast is. When testing for outliers, the results of four models were identified, "*Agriculture, Forestry, and Fishing*", "*Mining*", "*Electricity*", "*Hotels and Restaurants*".

When the same exercise was carried out for the MAPE statistics, three outliers were found, namely "*Mining*", "*Electricity*", and "*Hotels and Restaurant Activities*". Nevertheless, all except one model produced results which fell within the acceptable range of $\mu(Y_t) < 10\%$ as per Lewis (1982).

The Thiel U2 coefficient is a statistic that measures how well a model predicts changes in a variable. The closer the value is to 0, the better the forecast performance, however, if the value exceeds 1, the forecast is deemed inferior to a simple no change naïve forecast. From the industrial forecasts, seven were identified with a Thiel U2 coefficient exceeding 1, with three surpassing the acceptable range by large margins. These were "*Financial Activities*", "*Real Estate Activities*", and "*Other Service Activities*".

Incidentally, only the "*Mining*" model produced forecasts in which the RMSE and MAPE were outside the acceptable bounds, and their Thiel U2 coefficient surpassed the 1.0 acceptable range, albeit by a marginal level. This suggests that the model is poorly estimating the general level of the variable and also the direction and magnitude of changes. Thus, six forecasts failed to meet expectations in only one test, two failed in two tests, and only one failed in all three tests. Six of the forecasts produced excellent results, as identified by their RMSE, MAPE, and Thiel U2 coefficients. Notably, the worst performing models contributed the least to These results can be summarised in Table 8 below along with their contribution to GDP.

Table 8: Summary of Forecast Accuracy

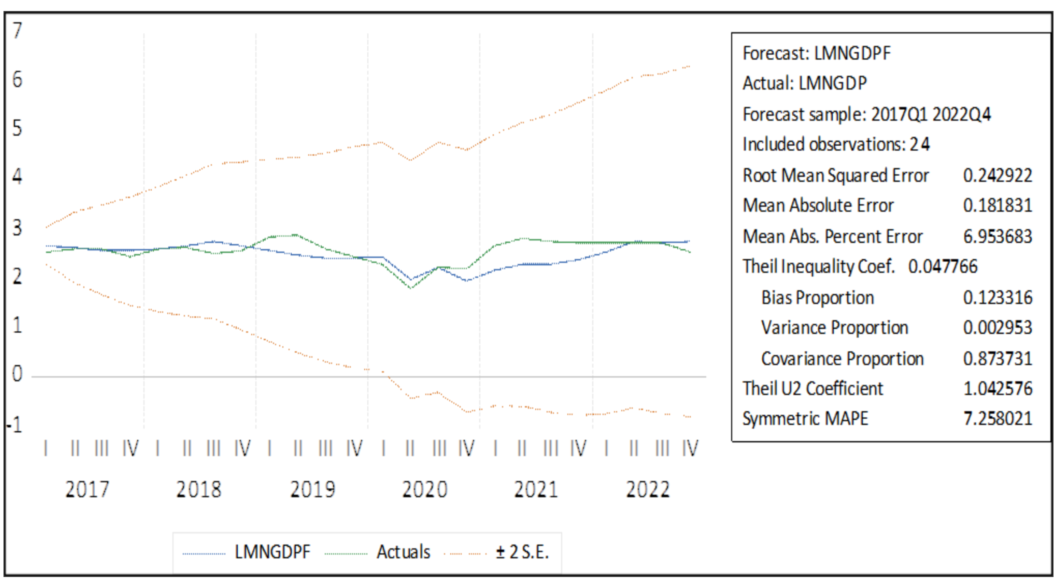
Number of Forecast Models	Number of Tests Failed	Contribution of Industries to GDP
1	3	2.3%
2	2	7.8%
6	1	46.7%
6	0	43.2%

5.0 Discussion

5.1 Industrial Model Performance

The worst performing model was "*Mining*", which did not forecast the general level or magnitude and direction of the dependent variable, the mining GVA. This can be attributed to a few reasons. The explanatory variable 'oil production value' was chosen as the better alternative indicator for use in the sectoral model, as financial data for crude oil extraction, oil exploration, and mining of minerals are not available on a high frequency basis. As was the case with tourist arrivals, the indicator was found to be highly correlated with the GDP category but not total GDP. Its inclusion in the model alongside the COVID2020 dummy variables, which served to capture the structural breaks following the pandemic, yielded weak results. A high MAPE highlighted the indicator's unpredictability in estimating movements in the category, which was reinforced when analysing the RMSE level as a percentage of the average value of the category (15.6%). Lastly, the Theil U2 coefficient indicated a poorly performing model. Nevertheless, given the category's average contribution to GDP (2.3%), and the lack of alternative indicators, the results were accepted. The full dynamic forecast results for the model are outlined in figure 4 below.

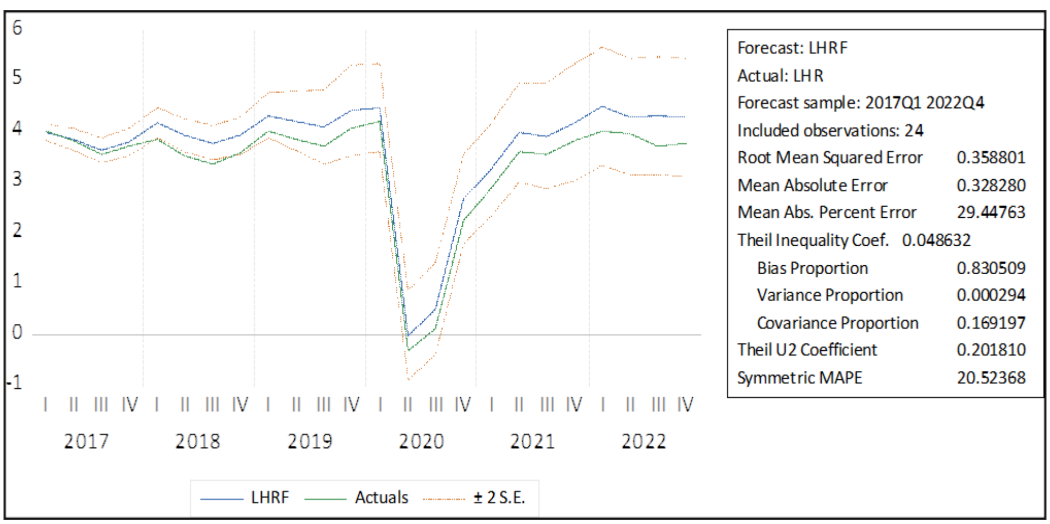
Figure 4: Dynamic Forecast Results for Mining Industrial Model



Of the two models which did not meet the requirements of two tests, "*Hotels and Restaurants*" performed the worst. Despite the model capturing the direction and magnitude of changes, it was unable to capture the general level of the variable. The index "*Tourist Arrival Index*" selected to nowcast "*Accommodation and Food Service Activities*" produced a high correlation coefficient when tested against the GVA. Thus, the indicator provides an accurate measure of activity in the GDP category.

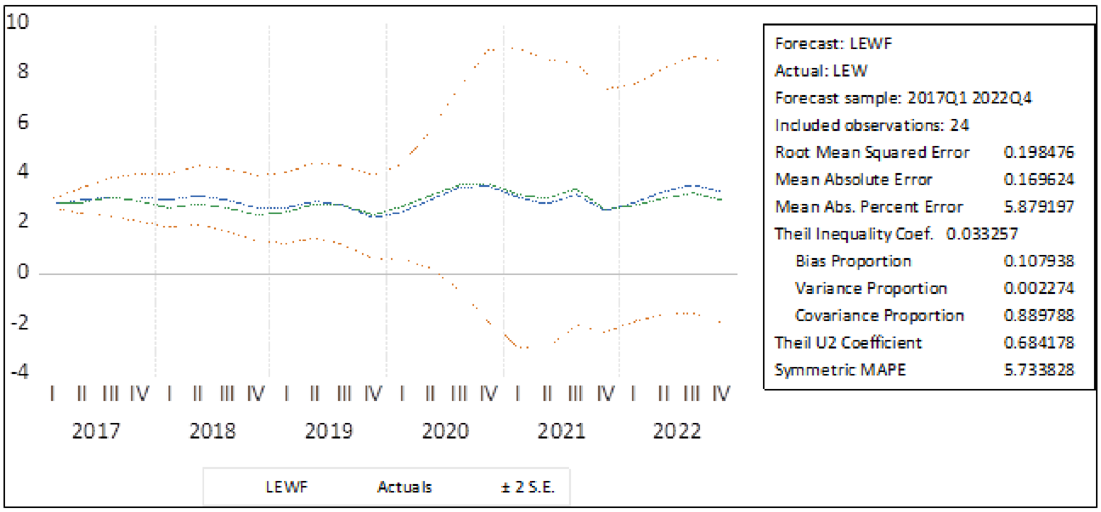
Nevertheless, as a substitute for actual expenditure on accommodation, food, and beverage services compiled annual by the SIB, the indicator did not sufficiently capture the dynamics of a given quarter, leading to the high RMSE/MAPE score. It is suspected that the dynamic nature of spending, particularly changes in average expenditure, cannot be accurately captured by arrival figures alone. Thus, this led to significant differences between actual and forecasted values, despite the model's capture of the general trend, as seen in figure 5 below. The inclusion of dummy variables for 2020 and helped to account for the structural break brought on by the event.

Figure 5: Dynamic Forecast Results for Hotel and Restaurants Industrial Model



By Lewis' (1982) metric, the forecast for "*Electricity*" had an acceptable MAPE score, below 10%. Nevertheless, when plotted, it proved to be an outlier, despite having an appropriate Thiel U2 coefficient. As was the case with "*Mining*", the lack of alternative variables led to only one explanatory variable, 'electricity generation' and the COVID2020 dummy to be used in the sectoral model. However, results were deemed acceptable and included in the overall calculation of total real GDP. Notably, In terms of average contribution to GDP between 2000 – 2022, "*Electricity*" was the smallest GDP category with a high MAPE. The results are presented in figure 6 below.

Figure 6: Dynamic Forecast Results for Electricity Industrial Model

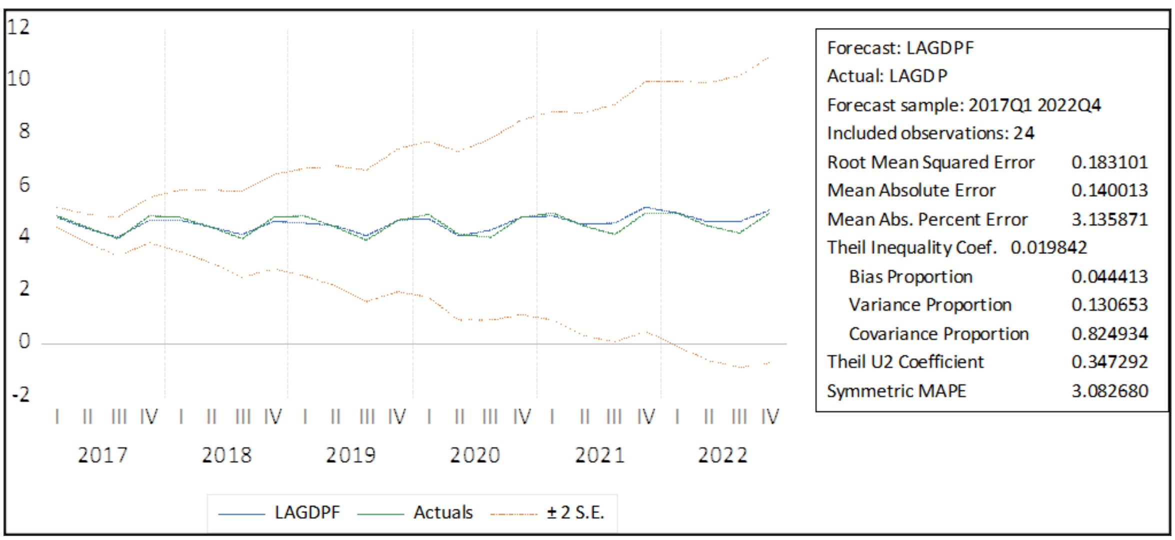


For the model "*Agriculture, Forestry and Fishing*" which produced acceptable MAPE and Thiel U2 coefficient, it was found that the RMSE value of 0.18, was the fourth highest. This performance could not be attributed to a specific factor, but rather a combination of multiple. The first is the composite index 'agricultural production index' which was used in the sectoral model to forecast "*Agriculture, Forestry, and Fishing*". The index comprised of 13 weighted indicators, representative of the 10 subindustries present within the GDP category. The index attained an acceptable correlation score of 0.63.

However, when observing the graphical representation of forecasted output versus actual output in figure 4.7 for "*Agriculture, Forestry, and Fishing*", the model underpredicts by wider margins than it overpredicts. The tendency to underpredict the gross value added of the category is suspected to result from the spike in real value for export commodities. The index "*LAG*" uses volume movements as a proxy for value movements which is used in the compilation of the GDP category, and thus, would not capture the price components. One alternative to this problem would be the use of alternative data.

As alluded to in the literature review, the use of alternative data would vastly improve the accuracy of nowcasting models if implemented correctly (Cascaldi-Garcia, Luciani, & Modugno, 2023). Attempting to improve the accuracy of their Agriculture GVA forecast, Kaustubh et al., (2024) incorporated rainfall data which may not be categorised strictly as alternative, but it does deviate from traditional variables as per the existing literature. The results of the forecast are illustrated in the figure 7 below.

Figure 7: Dynamic Forecast Results for Agriculture, Forestry, and Fishing Industrial Model



The remaining six models identified from Table 6 all registered Thiel U2 coefficients above 1.0. According to theory, this would make the forecast models unacceptable as they do not capture the direction and magnitude of changes. However, they all produced low RMSE and MAPE levels, indicating they estimate the forecasted GVA level to a sufficient degree. Additionally, only three of the models, namely "*Financial Activities*", "*Real Estate Activities*", and "*Other Service Activities*" register coefficients higher than the marginal 1.20. Thus, the acceptance of these results depends solely on the specific goals and priorities of this study.

This goal is defined as capturing the general level of the variable for aggregation to nowcast total GDP for a given quarter. With these considerations accounted for, the models are accepted solely based on these factors:

1. The total GDP forecast displays highly accurate results, this is expected given that the three models with high U2 coefficients only contribute to 20.5% of GDP.
2. The lack of timely indicators for these tertiary sector industries.
3. That revisions to these models will be forthcoming, either through an adjustment in the model structure, the inclusion of alternative data, or the incorporation of newly available exogenous variables.

5.2 Overview of Previous Model Framework

To demonstrate the improvement of the new framework of models, they are compared in the subsequent sections.

The previous nowcasting framework utilised a bottom-up sectoral approach to aggregate 11 sectoral estimates to produce a quarterly GDP forecast. Despite Arana's (2015) preference for indicator models, the bridge AR/MA approach was chosen from the options presented as it provided a more comprehensive explanation of the underlying drivers of economic growth. The models captured all the major subcategories of GDP, with all three major sectors (the primary, secondary, and tertiary sectors) accounted for.

5.2.1 Composition of Previous Models

The 11 bridge models were constructed as per Arana (2015), utilising a combination of OLS and ARMA/ARIMA models and bridge equations. Bridge equations are linear regression models that use high-frequency indicators to connect to a lower-frequency dependent variable. The model then uses the temporally aggregated values as regressors in the equation to obtain forecasts of the low-frequency variable.

Two models, "*Agriculture*" (*AGDP*) and "*Other Private Services*" (*OPS*), utilised OLS due to their unique variable parameters and followed the model:

$$Y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \ldots + \beta_p x_p + \epsilon_i \tag{1}$$

Where y_i is the dependent variable to be determined, β_0 is the intercept, $\beta_1 \ldots \beta_n$ is the coefficient, and $x_1 \ldots x_n$ are the independent variables.

The remainder of the sectoral models employed an ARMA approach, defined as:

$$Y_t + c + \phi_1 Y_{t-1} + \phi_2 T_{t-2} + \cdots + \phi_p T_{t-p} + \theta_1 T_{\epsilon t-1} + \theta_2 T_{\epsilon t-2} + \cdots + \theta_q T_{\epsilon t-q} + \epsilon_t \tag{2}$$

Where Y_t is the value of the time series at time t, the constant term is defined as c, $\phi_1, \phi_2 \ldots, \phi_p$ are the autoregressive coefficients at the order p, the moving average coefficients are $\theta_{1,}, \theta_2, \ldots, \theta_q$ at the order q, while ϵ_t represents the error term at time t.

5.2.2 Model Performance

Over the five-year period, following the implementation of the nowcasting framework, 21 quarterly forecasts were conducted. From the estimates, it was observed that the Mean Absolute Percentage Error (MAPE) stood at 6.1%, while the Root Mean Square Error (RMSE) equalled $47.2mn or 6.9% of the average actual GDP, and the Mean Absolute Deviation (MAD) totalled $8.1mn or 1.2% of the average actual GDP. Given its ease of interpretation, this section will focus on the MAPE value and errors. According to Lewis' (1982), MAPE results with values of less than 10.0% are defined as highly accurate forecasts. Thus, model outturns will be weighed against this metric.

As observed in Figure 8, discrepancies between the pre-rebasing GDP growth rate and the forecasted GDP growth rates became more pronounced in mid-2019, leading to the onset of the COVID-19 pandemic. Five of the 11 sectoral models displayed high MAPE values, as noted in Table 9, with the categories "Fishing," "Hotels and Restaurants," and "Producers of Government Services" recording consistent double-digit error values. Basic nowcasting principles stress the need for constant updates given the ever-changing relationship between indicators and GDP (Bragoli, Metelli, & Modugno, 2014). In sectoral forecasting, this deterioration of the explanatory power variables can be ascribed to idiosyncratic forces affecting the movement of small components of GDP (Dias, Pinhiero, & Rua, 2016). However, based on trend analysis and informed judgment, it was found that the underlying relationships between dependent and independent variables shifted for the five sectoral models, resulting in larger discrepancies.

Figure 8: Comparison of GDP Results for the Initial Nowcasting Framework

(Actual Vs Forecast)

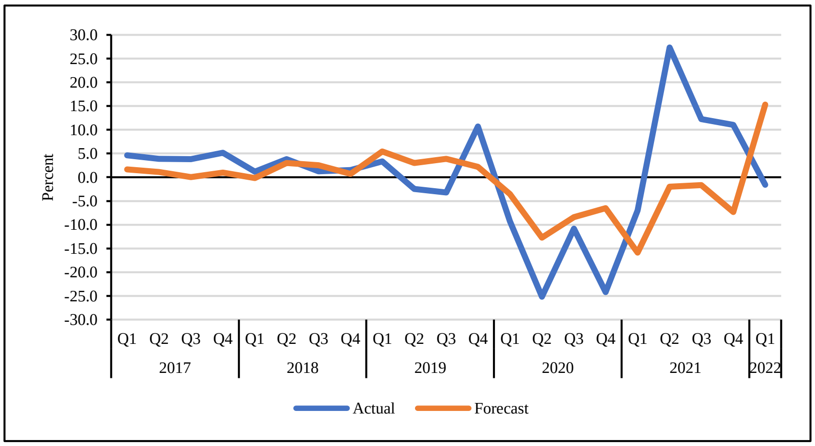


Table 9: Mean Absolute Percentage Error for Sectoral Models in Initial Nowcasting Framework (2017Q1 – 2022Q1)

	MAPE
Agriculture, hunting and forestry	5.7
Fishing	*29.0*
Manufacturing (incl. mining and quarrying)	7.2
Electricity and water	3.3
Construction	6.0
Wholesale and retail trade; repair	*9.5*
Hotels and restaurants	*14.9*
Transport and communication	*11.1*
Other private services, etc. FISIM 1)	2.6
Producers of government services	*15.0*
Taxes on products	4.2

5.3 Comparison of Quarterly GDP Results

To compare the results of the current and previous nowcasting frameworks, the descriptive metrics RMSE and MAPE were used to compare predicted and actual GDP estimates for the period 2017Q1 and 2022Q1[4].

Table 10: Comparison of Real GDP Performance Metrics Between Previous and Current Nowcasting Frameworks

	Old Nowcasting Framework	New Nowcasting Framework
Root Mean Squared Error (RMSE)	47.2	28.5
Mean Absolute Percentage Error (MAPE)	6.1	2.1

For Real GDP between 2017Q1-2022Q1, the new nowcasting framework outperformed the old in the two metrics employed. Further insight into the performance of the new framework was provided from the analysis of the RMSE as a percentage of average quarterly GDP, which increased by a significant 60.0% margin during the rebasing exercise. As a result, the RMSE as a percentage of average GDP improved by 4.6 percentage points from 6.9% to 2.6%.

Table 11: Comparison of MAPE between the Pre- and Post-Pandemic Era

	Old Nowcasting Framework	New Nowcasting Framework
Pre-Pandemic Era (2017Q1 - 2019Q4)	4.1	1.7
Post-Pandemic Era (2020Q1 - 2022Q1)	8.5	2.1

When comparing both models' MAPE performances in both the pre- and post-pandemic era, the new nowcasting framework significantly outperforms the old, as illustrated in Table 11 above. For the pre-pandemic period, the nowcasting framework's RMSE stood at 22.7 or 2.0% of average quarterly GDP, while the previous framework recorded 36.1, equivalent to 5.2% of average quarterly

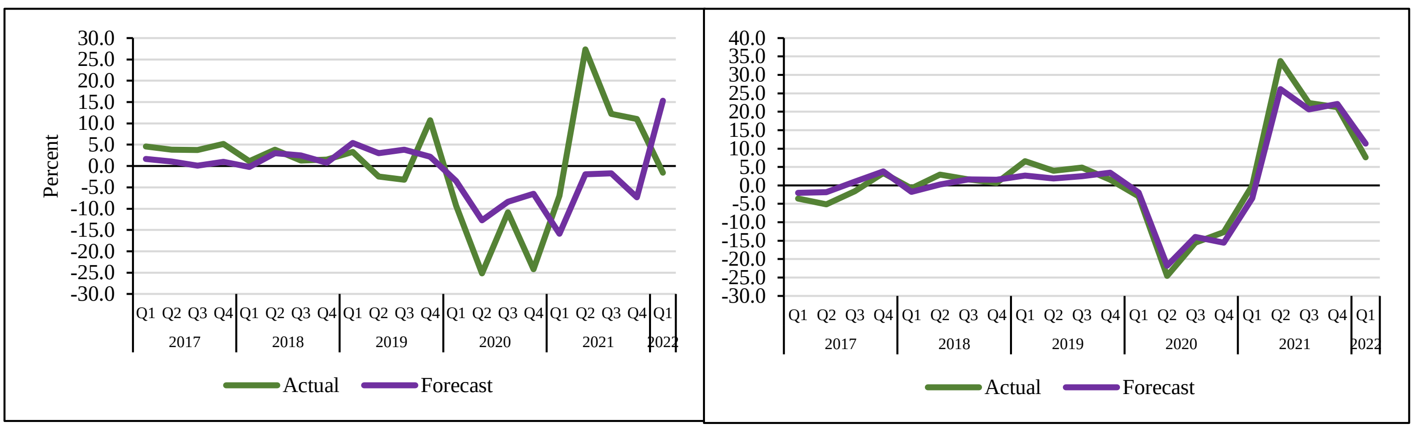
[4] Despite estimates generated for 2017Q1 – 2022Q4 under the new framework, to compare like=to-like figures from the old framework, the sample was reduced to 2017Q1-2022Q1.

GDP. The general improvement in predictive performance is attributable to the tightening of sectoral models' parameters, and the expansion in variables utilised.

In the post pandemic era (2020Q1-2022Q1), with a MAPE of 1.9%, the new model outperformed the previous model's 8.5%. Notably, in the previous model, the MAPE value more than doubled between the pre- and post-pandemic estimates, highlighting the models' inability to accurately respond to severe multi-sectoral shocks such as the COVID-19 pandemic.

The incorporation of the dummy variables capturing both the contraction period (2020Q1-2020Q4) and rebound period (2021Q1-2022Q1) across 13 of the 16 models drastically improved results as illustrated in Figure 9.

Figure 9: Previous vs New Framework's Comparative Percentage Change in GDP (Actual Vs Forecast)



5.4 Major Takeaways

This paper's aim was twofold; enhance the Bank's GDP nowcasting framework to accurately capture the recent structural changes brought on by the 2022 rebasing exercise and to effectively capture the impact of economic shocks. To achieve this, the existing nowcasting framework was expanded to a total of 16 sectoral models utilising 45 indicators, inclusive of 2 dummy variables. As per the bottom-up approach, the estimates produced by the models were aggregated to arrive at a GDP estimate. Across the forecast period, the framework produced an accurate estimate of GDP, significantly bolstering the predictive power of the previous framework as proven by the RMSE and

MAPE values tested. Nevertheless, critical observations were made regarding the complexity of maintaining the current number of models, as well as constraints brought on by the lack of timely indicators with sufficient explanatory power.

With all 16 sectoral models estimating their respective GDP categories and performing at variable levels of accuracy, the likelihood for issues arising within a given model from quarter-to-quarter warrants pause. Evidence for this arose from the observation of varying MAPE levels for the models, which highlighted four models with higher-than-average levels of inaccuracy. Notably, the three worst performing models were tied to three of the smallest GDP categories in terms of their contribution to GDP. It was suggested in one study that the smaller components of GDP behave in a more volatile manner, driven by idiosyncratic forces (Dias, Pinhiero, & Rua, 2016). Alternatively, it is outlined in this paper that the weakness may stem from the explanatory variables chosen. Nevertheless, the models were kept in the framework despite their performance, as their contribution to the overall GDP estimate is insignificant.

Another pertinent downside to maintaining a high number of sectoral models, was a higher likelihood of the loss of predictive power by an indicator over a period of time. This can manifest particularly in any of the indices created with weights assigned to each indicator. This was surmised to be the case within the "*Professional and Administrative Services*" model, in which the BPO activity indicator grew in terms of its contribution to growth in the GDP category, diminishing the statistical importance of other indicators present in the index. This could be addressed in a number of ways, with the most likely outcome entailing the elimination of statistically insignificant variables within indices and/or the inclusion of new variables.

The complement of variables included in the nowcasting framework were found to accurately predict the movement of quarterly GDP with the exception of a few that were either poorly correlated or did not provide sufficient causation. Nevertheless, multiple indicators presented structural breaks brought on by the pandemic. This was rectified within the models by the inclusion of two dummy variables capturing either the contraction, the rebound, or both, across 13 of the models following the pandemic which allowed for improvements in model fitness. As a result, it was determined that the general accuracy of the sectoral models warranted their explicit inclusion and expansion, as they may be able to provide critical insight to policy makers monitoring targeted industries.

6.0 Conclusion

An accurate, reliable, and timely nowcasting estimator of quarterly GDP was deemed necessary to grasp the movements within the economy following the pandemic. This paper went a step further and reinforced the use of sectoral forecasts to provide policymakers with an added dimension of information to better inform their decisions. Altogether, despite the challenges in implementing a framework of this scale, the trade-off provided accurate estimates of not only quarterly GDP on a whole, but sectoral movements present within the various categories of GDP.

It is critical to note that similarly to the way in which this framework was born out of a need following the pandemic and revised to reflect the GDP rebasing exercise, further revisions will be necessary as Belize's economy matures. With the availability of a complement of sophisticated techniques aimed at estimating quarterly GDP, and the race towards producing results with the least amount of lag time, room exists to improve the framework considerably. As explored by Nakazawa (2022), non-traditional indicators may be implemented to supplement underperforming models ran utilising less-than-ideal variables.

The decision to link the current framework to the statistical offices' production-based approach to calculating GDP prevented the inclusion of alternative data and restricted the number of variables utilised in this study. Thus, the impact of external variables such as the nation's leading trade partner's GDP growth, inflation, and price indices played no role in the generation of the current iteration of the framework.

A worthwhile effort would be the exploration and use of alternative variables simultaneously to generate a hybridised indicator list. This could prove to be a steppingstone to build a far more robust framework. It is only through the understanding that the only constant in nowcasting is change, will more robust frameworks be produced. In doing so, Belize's position on remaining at the forefront of the nowcasting cutting edge.

8.0 References

Akpan, P. E., & Udofia, L. E. (2017). Economics Methodology and The Financial Crisis: A RE-Examination. *The International Journal of Humanities and Social Studies, 5*(4). Retrieved from https://www.internationaljournalcorner.com/index.php/theijhss/article/view/125315

Arana, R. (2015). *Evaluating Short Term Bridge Model Forecasts of Real GDP in Belize: A Quarterly Approach*. Central Bank of Belize. Retrieved from https://www.centralbank.org.bz/docs/default-source/4.5-conferences-and-working-papers/evaluating-short-term-bridge-model-forecasts-of-real-gdp-in-belize---a-quarterly-approach.pdf?sfvrsn=5d908e35_0

Barrios, J., Martin , L., Escobar, J., Pena, W., & Leslie , J. (2021). *Nowcasting to Predict Economic Activity in Real Time: The Case of Belize and El Salvador*. Retrieved from Inter-American Development Bank.

Bragoli, D., Metelli, L., & Modugno, M. (2014). *The Importance of Updating: Evidence from a Brazilian Nowcasting Model.* Federal Reserve Board, Divisions of Research and Statistics and Monetary Affairs, Washington D.C. doi:https://doi.org/10.17016/FEDS.2014.94

Carriero, A., Clark, T., Marcellino, M., & Mertens, E. (2024). Addressing COVID-19 Outliers in BVARs with Stochastic Volatility. *The Review of Economics and Statistics, 106*(5), 1403-1417. doi:https://doi.org/10.1162/rest_a_01213

Cascaldi-Garcia, D., Luciani, M., & Modugno, M. (2023). *Lessons from Nowcasting GDP across the World.* Washington: Board of Governors of the Federal Reserve System. International Finance Discissions Papers 1385. doi:https://doi.org/10.17016/IFDP.2023.1385

Dauphin, J.-F., Dybczak, K., Maneely, M., Sanjani, M., Suphaphiphat, N., Wang, Y., & Zhang, H. (2022). *A Scalable Approach Using DFM, Machine Learning and Novel Data, Applied to European Economies*. IMF Working Paper 22/52, International Monetary Fund, European Department.

Dembure, H., & Ziramba, E. (2016). Testing the Validity of Wagner's Law in the Namibian Context: A Toda-Yamamoto (TY) Granger Causality Approach, 1991 - 2013. *The Journal of Economic Policy Reform*, 19(2) 53-70. doi:https://doi.org/10.1080/17487870.2015.1114132

Dias, F., Pinhiero, M., & Rua, A. (2016). A Bottom-Up Approach for Forecasting GDP in a Data Rich Environment. *Applied Economics Letters, 25*(10), 718-723. doi:https://doi.org/10.1080/13504851.2017.1361000

Foroni, C., Marcellino, M., & Stevanovic, D. (2022). Forecasting the Covid-19 Recession and Recovery: Lessons from the Financial Crisis. *International Journal of Forecasting*, 596-612. doi:https://doi.org/10.1016/j.ijforecast.2020.12.005

Furceri, D., Loungani, P., Ostry, J. D., & Pizzuto, P. (2021). Will COVID-19 Have Long-Lasting Effects on Inequality? Evidence from the Past Pandemics. *IMF Working Paper, 21, 127*. International Monetary Fund.

Giannone, D., Reichlin, L., & Small, D. H. (2006). Nowcasting GDP and Inflation The Real-Time Informational Content of Macroeconomic Data Releases. *ECB Working Paper Series(No. 633)*.

Hahn, E., & Skudelny, F. (2008). Early Estimates of Euro Area Real GDP Growth: A Bottom Up Approach from the Production Side. *ECB Working Paper Series(No. 975)*.

Kaustubh, N. A., Bhadury, S. S., & Ghosh, S. (2024). Reinvigorating GVA Nowcasting in the Post-Pandemic Period: A Case Study for India. *Bulletin of Monetary Economics and Banking, Special Issue*, 95-130. doi:https://doi.org/10.59091/2460-9196.2160

Lewis, C. D. (1982). International and Business Forecasting Methods. *Journal Of Forecasting, 2*(2). doi: https://doi.org/10.1002/for.3980020210

Maccarrone, G., Morelli, G., & Spadaccini, S. (2021). GDP Forecasting: Machine Learning, Linear or Autoregression? *Frontiers in Artificial Intelligence, 4*(757864). doi:10.3389/frai.2021.757864

Nakazawa, T. (2022). Constructing GDP Nowcasting Models Using Alternative Data. *Bank of Japan Working Paper Series, No. 22-E-9*.

Sankaran, A., Kumar, S., Arjun, K., & Das, M. (2019). Estimating the Causal Relationship Between Electricity Consumption and Industrial Output: ARDL bounds and Toda-Yamamoto Approaches for Ten Late Industrialised Countries. *Heliyon, 2405-8440*. doi:https://doi.org/10.1016/j.heliyon.2019.e01904

SIB. (2020). *Belize System of National Accounts Rebasing to Reference Year 2014: Results.* Economic Statistics, Statistical Institute of Belize. Retrieved from https://sib.org.bz/wp-content/uploads/20220718_BSNA_RebasingDocumentation.pdf

Uttam Chandra, P. (2020). The Causal Relationship Between Private Sector Credit Growth and Economic Growth in Bangladesh: Use of Toda-Yamamoto Granger Causality Test in VAR Model. *Munich Personal RePEc Archive(104476)*. Retrieved from https://mpra.ub.uni-muenchen.de/104476/

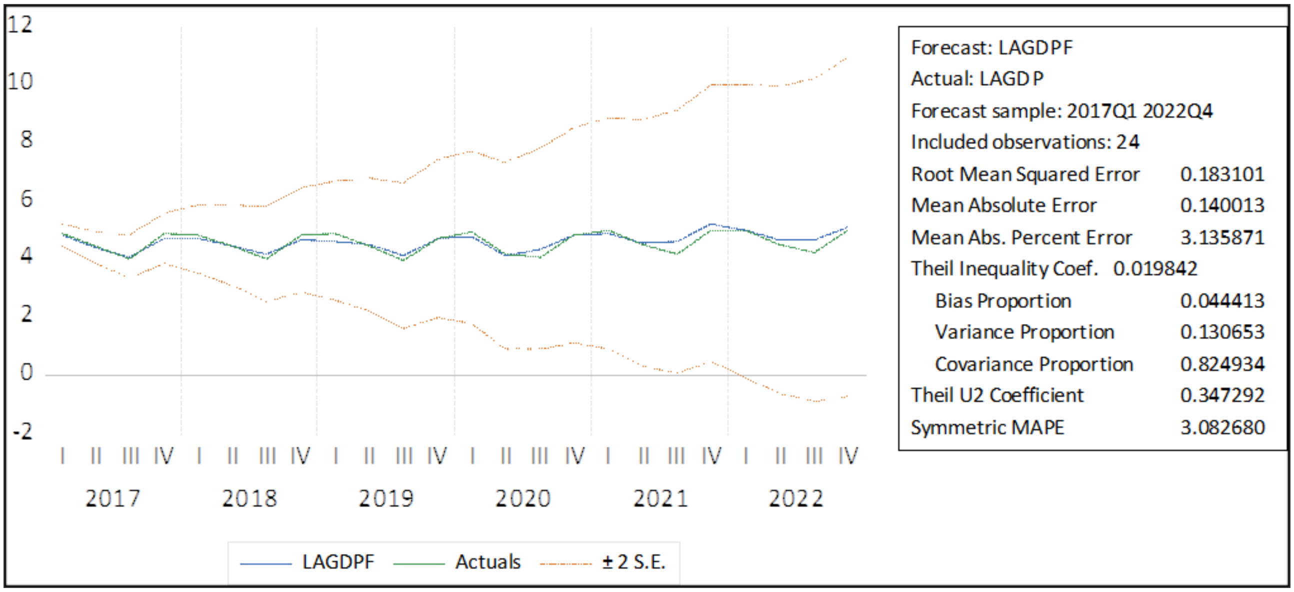
9.0 Appendix

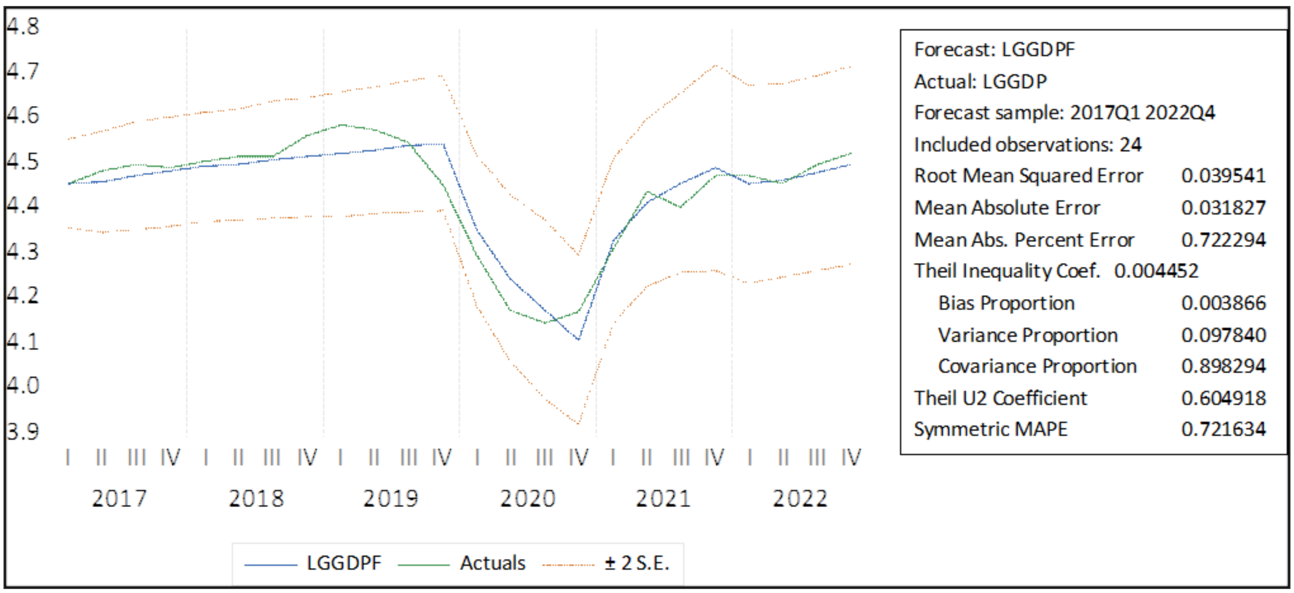
Appendix Table 1: Reclassification of Industrial Groupings

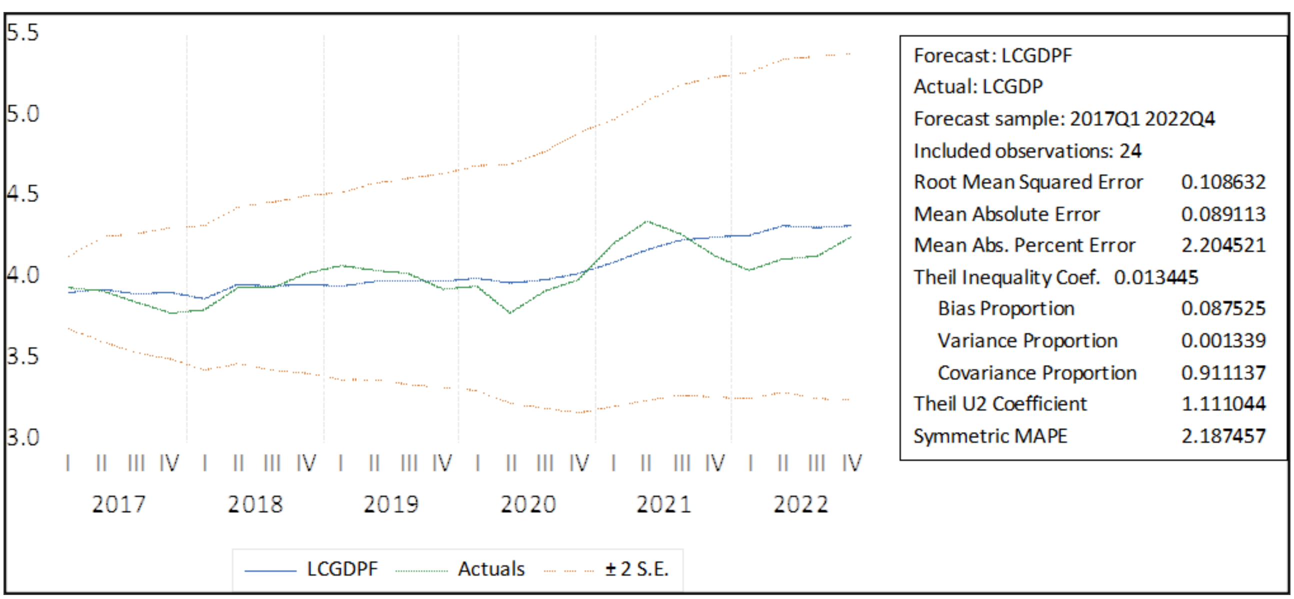
Base Year 2000	**Base Year 2014**	**Notes**
Agriculture, hunting and forestry	Agriculture, forestry and fishing	Absorbed "Fishing"
Fishing	Mining	Disaggregated from "Manufacturing"
Manufacturing (incl. mining and quarrying)	Manufacturing	
Electricity and water	Electricity	Disaggregated from "Electricity and Water"
	Water supply	Disaggregated from "Electricity and Water"
Construction	Construction	
Wholesale and retail trade; repair	Wholesale and retail trade	
Hotels and restaurants	Accommodation and food service activities	
Transport and communication	Transportation	Disaggregated from "Transport and Communication"
	Information and communication	Disaggregated from "Transport and Communication"
Other private services exc. FISIM 1)	Financial and insurance activities	Disaggregated from "Other Private Services"
	Real estate activities	Disaggregated from "Other Private Services"
	Professional scientific and technical activities	Disaggregated from "Other Private Services"
	Adminstrative and support service activities	Disaggregated from "Other Private Services"
	Arts, entertainment, recreation	Disaggregated from "Other Private Services"
	Other service activities	Disaggregated from "Other Private Services"
	Activities of households as employers	Disaggregated from "Other Private Services"
Producers of government services	Public administration and defence	Disaggregated from "Producers of Government Services"
	Education	Disaggregated from "Producers of Government Services"
	Human health and social work activities	Disaggregated from "Producers of Government Services"
Taxes on products	Taxes and Subidies	

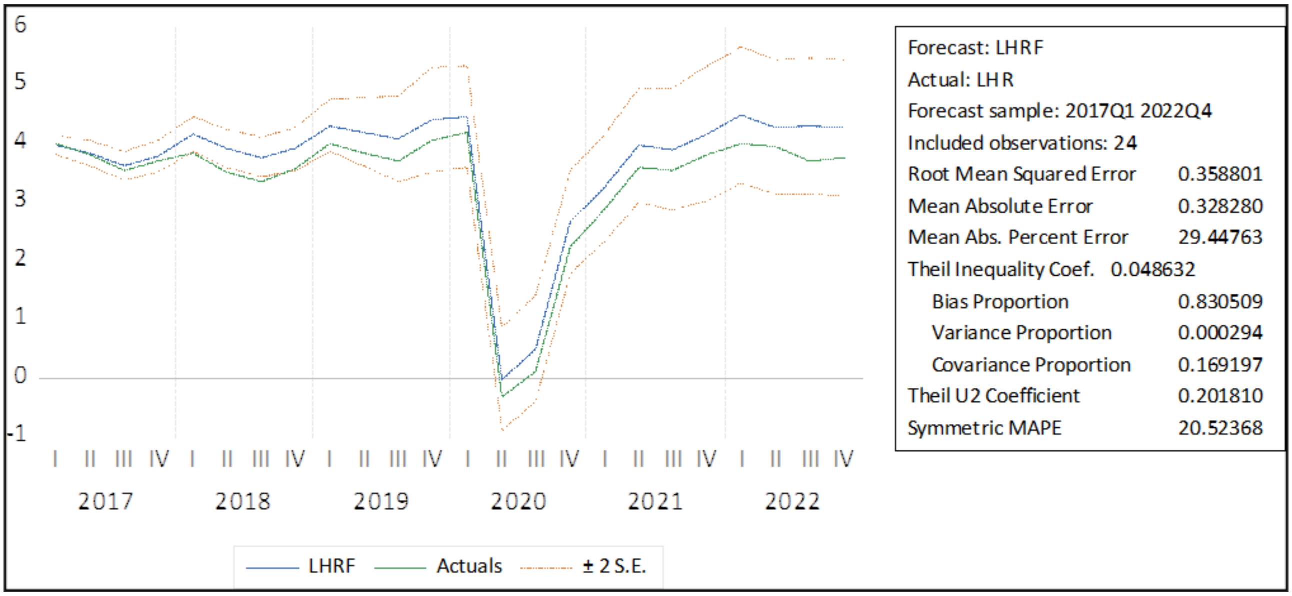
Appendix Table 2: Indicator and Index Component List

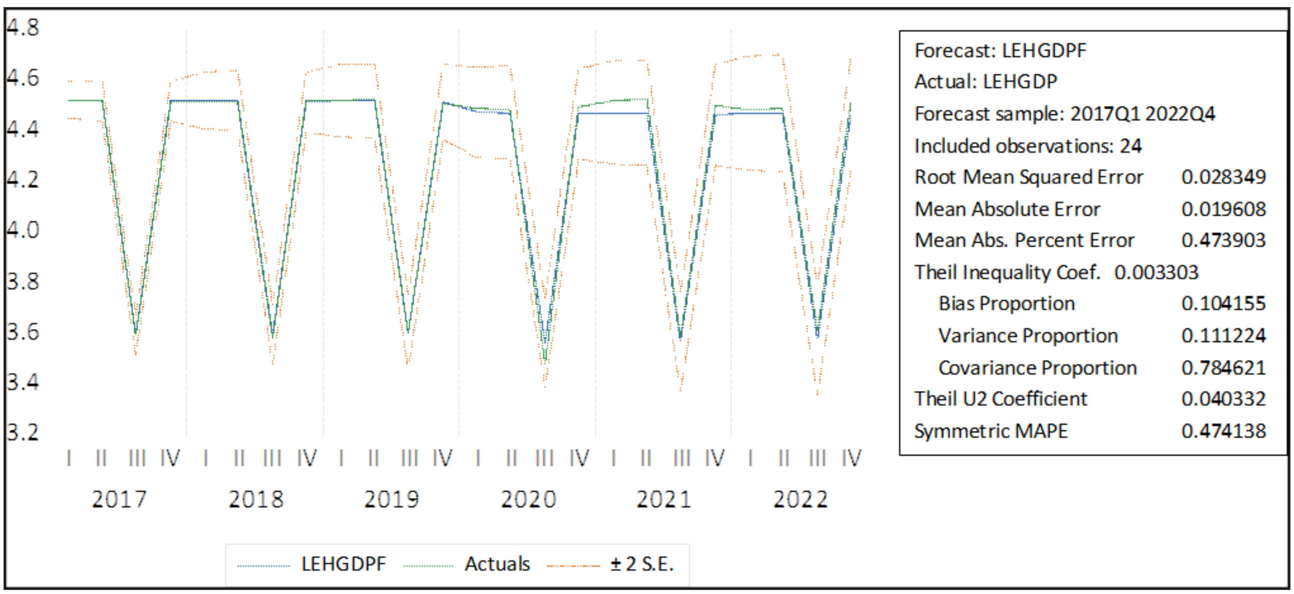
Sector	Index	Indicator	Frequency
Primary	**Agricultural Index**	Grapefruit Deliveries	Monthly
		Orange Deliveries	Monthly
		Banana Production	Monthly
		Sugarcane Production	Monthly
		Other Crops	Quarterly
		Chicken	Quarterly
		Beef	Quarterly
		Conch	Monthly
		Farmed Shrimp	Monthly
		Fish	Monthly
		Lobster	Monthly
		Oil Production Value	Monthly
Secondary	**Manufacturing Index**	Beer	Quarterly
		Flour	Quarterly
		Soft Drinks	Quarterly
		ONFC	Monthly
		GNFC	Monthly
		Grapefruit Concentrates	Monthly
		Orange Concentrates	Monthly
		Animal Feed	Monthly
		Fertilizer	Quarterly
		Dairy Production	Quarterly
		Sugar Production	Monthly
		Electricity Production	Monthly
		Water Distribution	Monthly
		Cement Imports	Quarterly
		Loans for Construction	Monthly
Tertiary & Taxes	**Tourist Arrival Index**	Air Arrivals	Monthly
		Sea Arrivals	Monthly
		Cruise Arrivals	Monthly
		Land Arrivals	Monthly
	Transportation Index	Sugar Production	Monthly
		Total Tourist Arrivals	Monthly
		Gross Merchandise Imports	Monthly
	Professional and Administrative Services Index	Business Tax	Monthly
		Business Processing and Outsourcing Income	Monthly
		General Sales Tax Collections	Monthly
		Government Current Expenditure	Monthly
		Government Wages and Salaries	Monthly
		Government Current Revenue	Monthly
		Total Loans	Monthly
		Money Supply	Monthly
		Saria	Monthly
		Gross Merchandise Imports	Monthly
		Total Tourist Arrivals	Monthly
		General Sales Tax Collections	Monthly

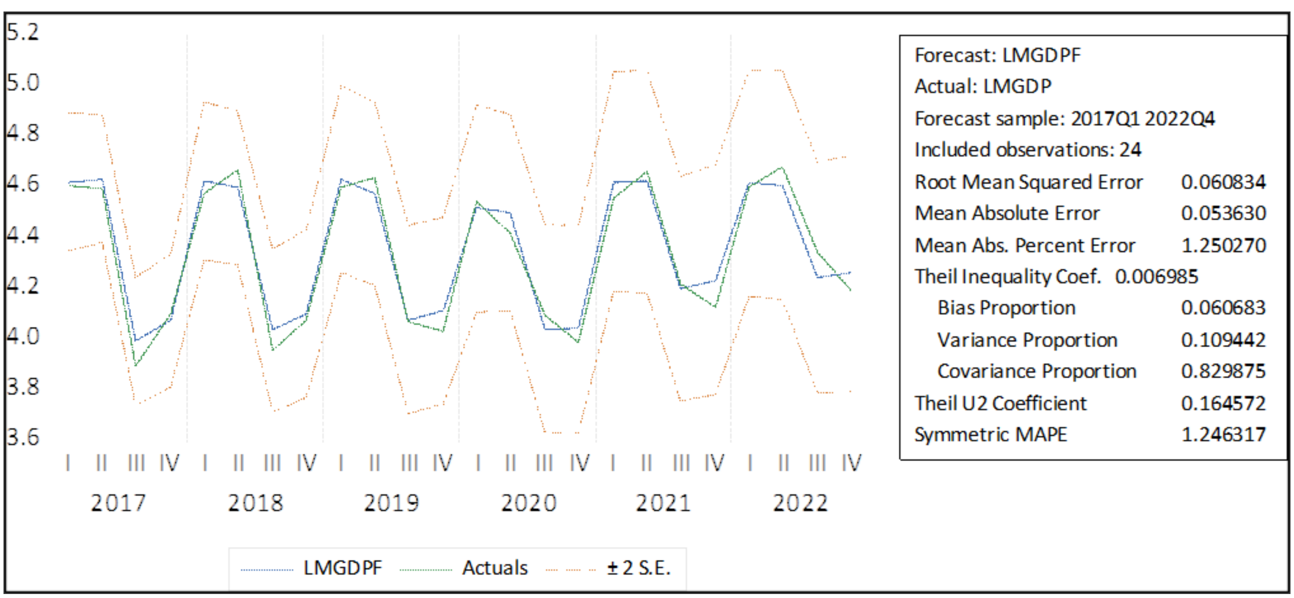

Forecast: LAGDPF
Actual: LAGDP
Forecast sample: 2017Q1 2022Q4
Included observations: 24
Root Mean Squared Error 0.183101
Mean Absolute Error 0.140013
Mean Abs. Percent Error 3.135871
Theil Inequality Coef. 0.019842
Bias Proportion 0.044413
Variance Proportion 0.130653
Covariance Proportion 0.824934
Theil U2 Coefficient 0.347292
Symmetric MAPE 3.082680
2017
2018
2019
2020
2021
2022
LAGDPF
Actuals
± 2 S.E.

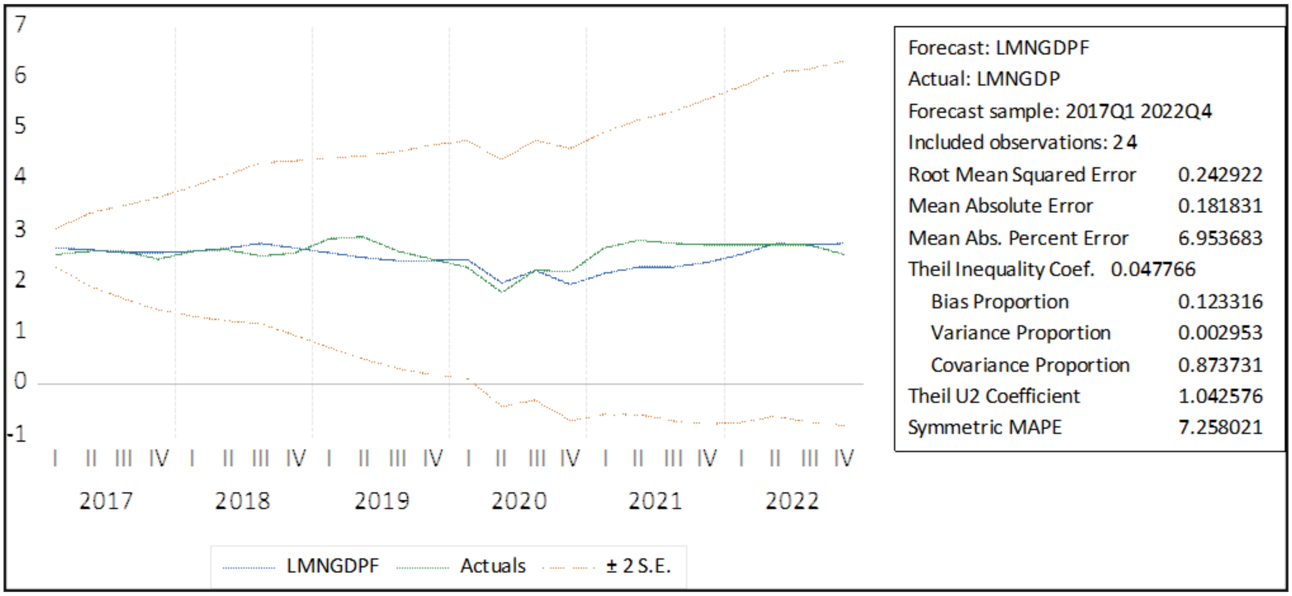

Forecast: LGGDPF
Actual: LGGDP
Forecast sample: 2017Q1 2022Q4
Included observations: 24
Root Mean Squared Error 0.039541
Mean Absolute Error 0.031827
Mean Abs. Percent Error 0.722294
Theil Inequality Coef. 0.004452
Bias Proportion 0.003866
Variance Proportion 0.097840
Covariance Proportion 0.898294
Theil U2 Coefficient 0.604918
Symmetric MAPE 0.721634
2017
2018
2019
2020
2021
2022
LGGDPF
Actuals
± 2 S.E.

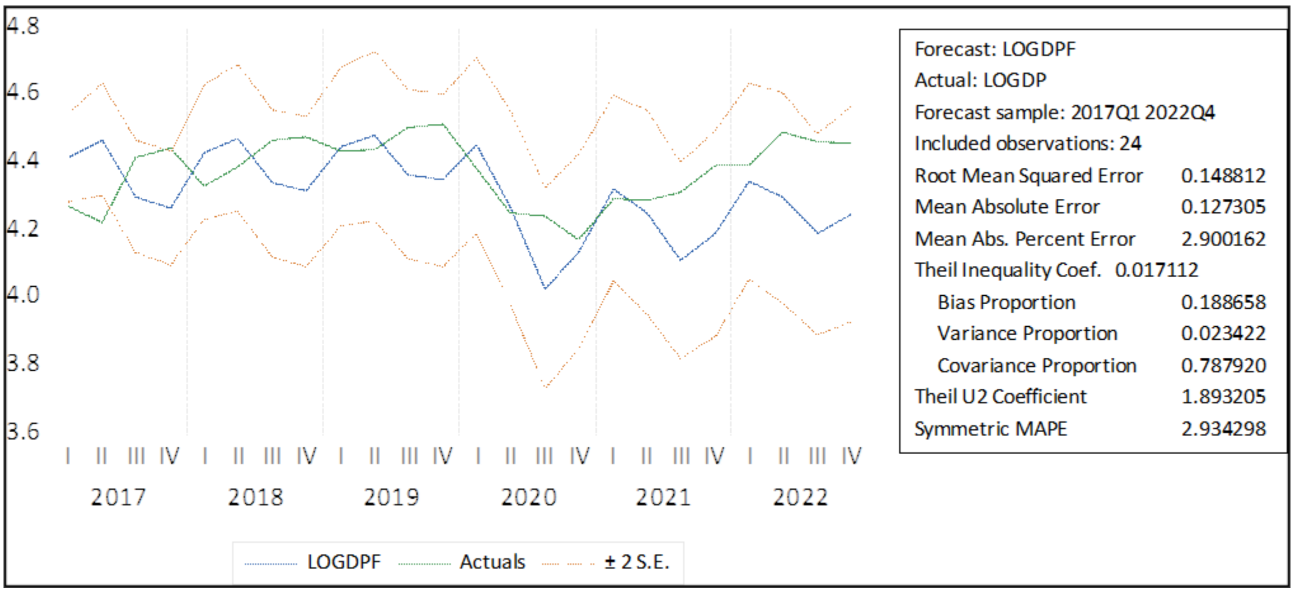

Forecast: LCGDPF
Actual: LCGDP
Forecast sample: 2017Q1 2022Q4
Included observations: 24
Root Mean Squared Error 0.108632
Mean Absolute Error 0.089113
Mean Abs. Percent Error 2.204521
Theil Inequality Coef. 0.013445
Bias Proportion 0.087525
Variance Proportion 0.001339
Covariance Proportion 0.911137
Theil U2 Coefficient 1.111044
Symmetric MAPE 2.187457
2017
2018
2019
2020
2021
2022
LCGDPF
Actuals
± 2 S.E.

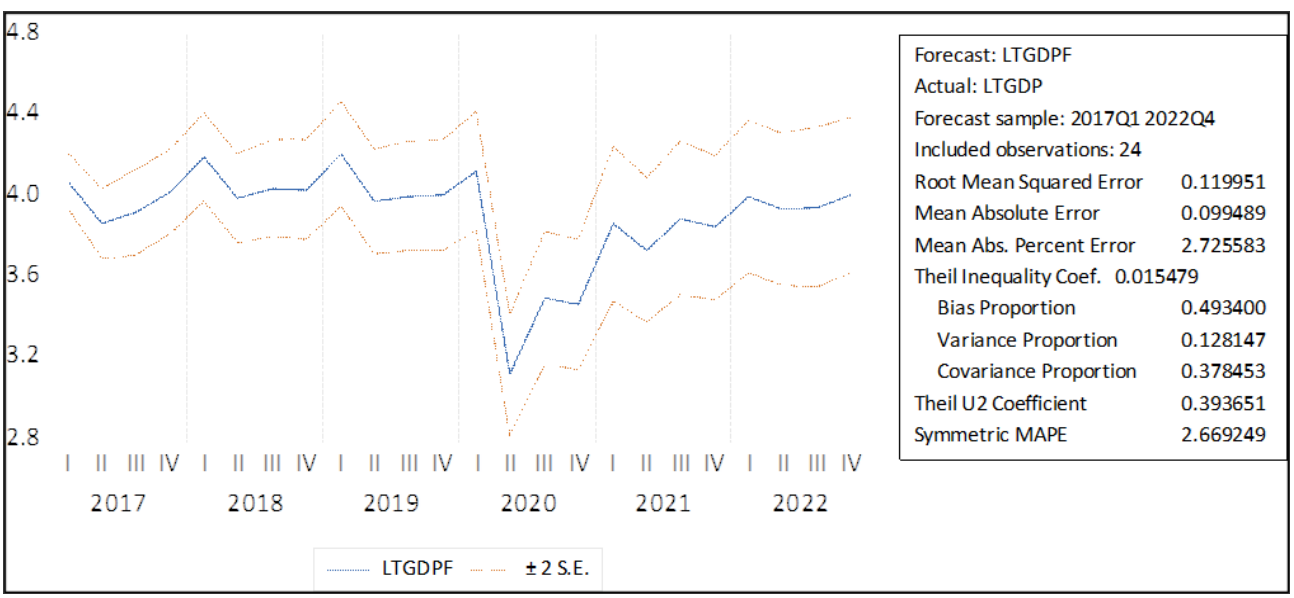

Forecast: LHRF
Actual: LHR
Forecast sample: 2017Q1 2022Q4
Included observations: 24
Root Mean Squared Error 0.358801
Mean Absolute Error 0.328280
Mean Abs. Percent Error 29.44763
Theil Inequality Coef. 0.048632
Bias Proportion 0.830509
Variance Proportion 0.000294
Covariance Proportion 0.169197
Theil U2 Coefficient 0.201810
Symmetric MAPE 20.52368
2017 2018 2019 2020 2021 2022
LHRF Actuals ± 2 S.E.

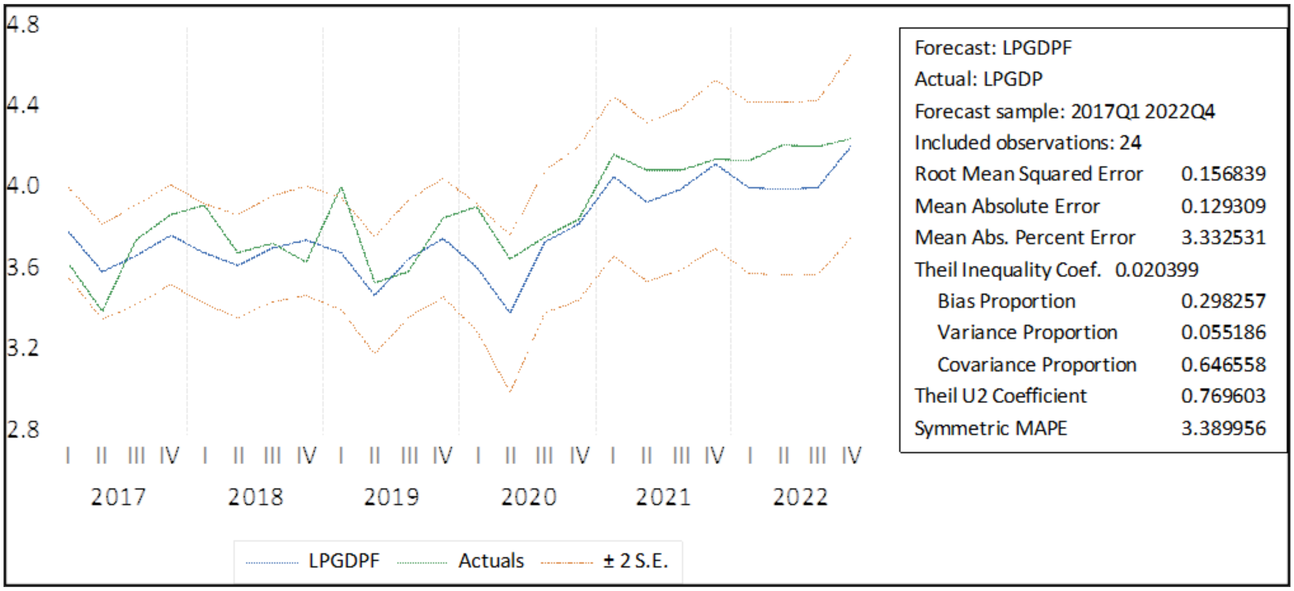

Forecast: LEHGDPF
Actual: LEHGDP
Forecast sample: 2017Q1 2022Q4
Included observations: 24
Root Mean Squared Error 0.028349
Mean Absolute Error 0.019608
Mean Abs. Percent Error 0.473903
Theil Inequality Coef. 0.003303
Bias Proportion 0.104155
Variance Proportion 0.111224
Covariance Proportion 0.784621
Theil U2 Coefficient 0.040332
Symmetric MAPE 0.474138
2017 2018 2019 2020 2021 2022
LEHGDPF Actuals ± 2 S.E.

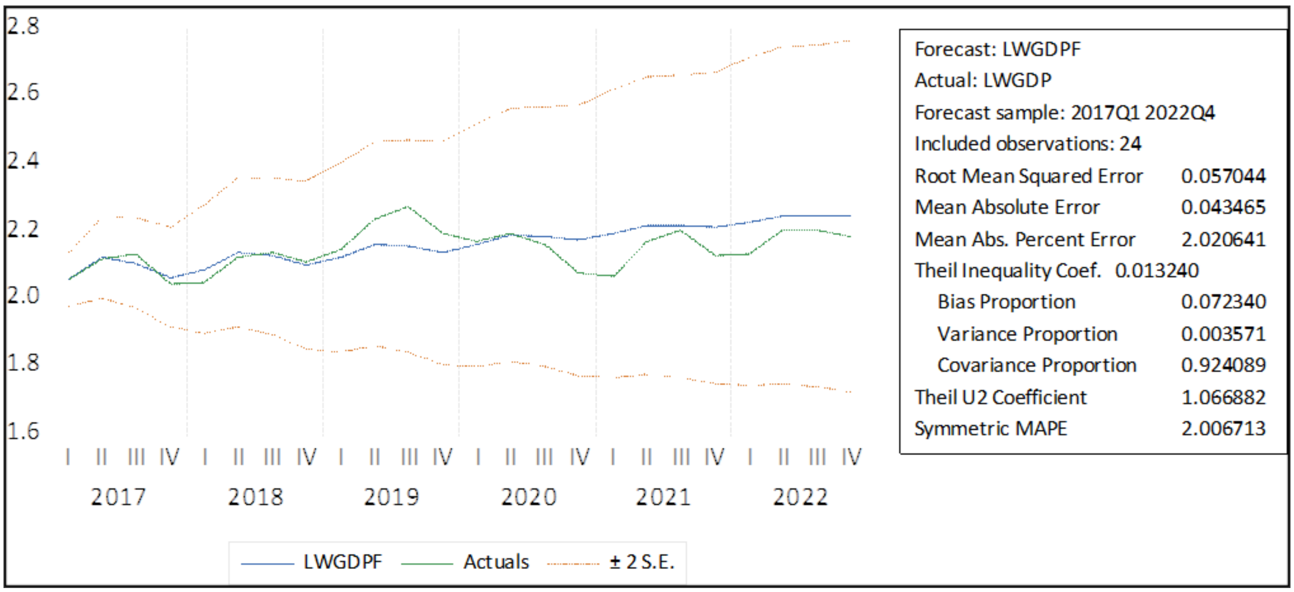

Forecast: LMGDPF
Actual: LMGDP
Forecast sample: 2017Q1 2022Q4
Included observations: 24
Root Mean Squared Error 0.060834
Mean Absolute Error 0.053630
Mean Abs. Percent Error 1.250270
Theil Inequality Coef. 0.006985
Bias Proportion 0.060683
Variance Proportion 0.109442
Covariance Proportion 0.829875
Theil U2 Coefficient 0.164572
Symmetric MAPE 1.246317
2017 2018 2019 2020 2021 2022
LMGDPF Actuals ± 2 S.E.

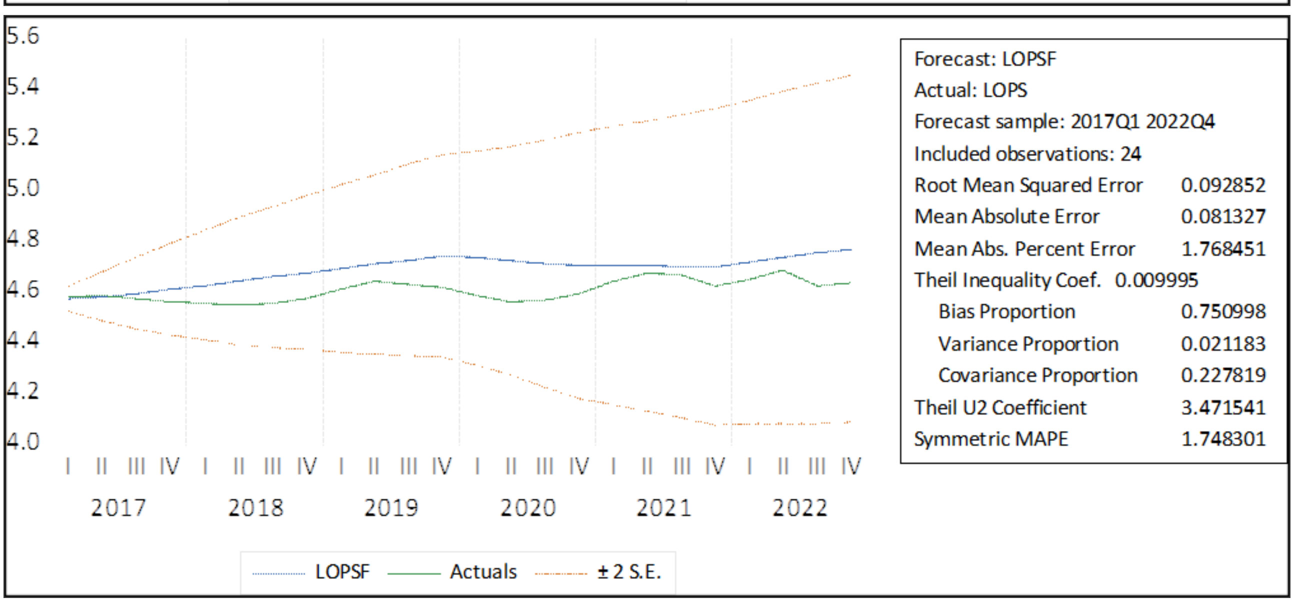

7
6
5
4
3
2
1
0
-1
I II III IV I II III IV I II III IV I II III IV I II III IV I II III IV
2017 2018 2019 2020 2021 2022
LMNGDPF
Actuals
± 2 S.E.
Forecast: LMNGDPF
Actual: LMNGDP
Forecast sample: 2017Q1 2022Q4
Included observations: 24
Root Mean Squared Error 0.242922
Mean Absolute Error 0.181831
Mean Abs. Percent Error 6.953683
Theil Inequality Coef. 0.047766
Bias Proportion 0.123316
Variance Proportion 0.002953
Covariance Proportion 0.873731
Theil U2 Coefficient 1.042576
Symmetric MAPE 7.258021

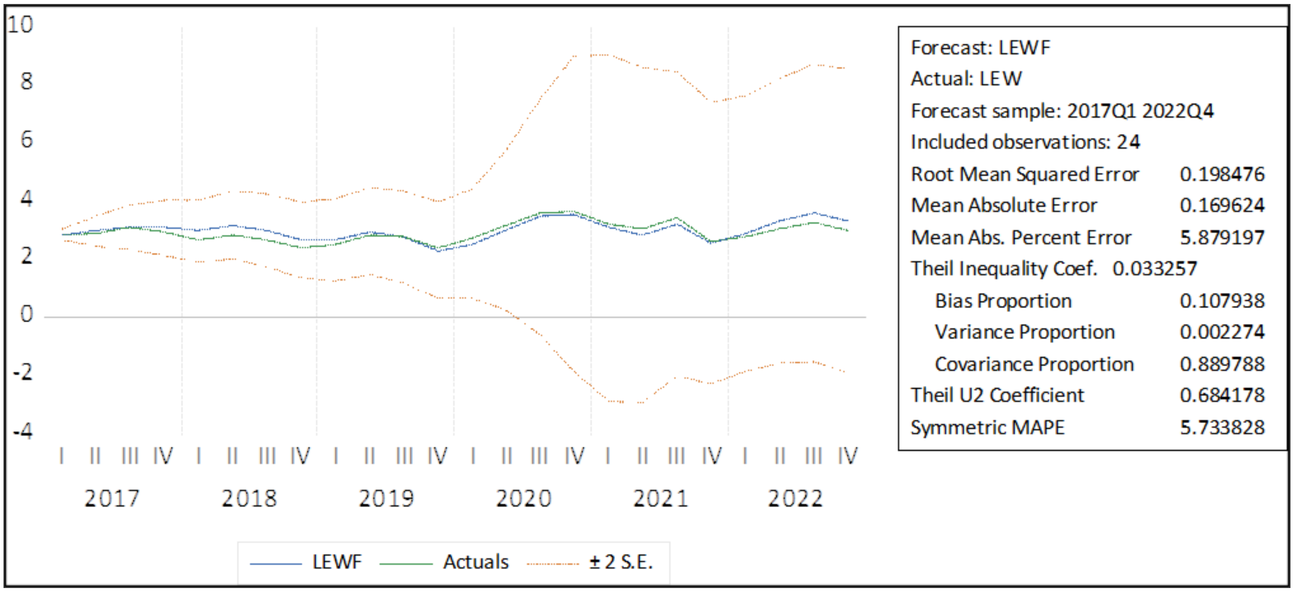

4.8
4.6
4.4
4.2
4.0
3.8
3.6
I II III IV I II III IV I II III IV I II III IV I II III IV I II III IV
2017 2018 2019 2020 2021 2022
LOGDPF
Actuals
± 2 S.E.
Forecast: LOGDPF
Actual: LOGDP
Forecast sample: 2017Q1 2022Q4
Included observations: 24
Root Mean Squared Error 0.148812
Mean Absolute Error 0.127305
Mean Abs. Percent Error 2.900162
Theil Inequality Coef. 0.017112
Bias Proportion 0.188658
Variance Proportion 0.023422
Covariance Proportion 0.787920
Theil U2 Coefficient 1.893205
Symmetric MAPE 2.934298

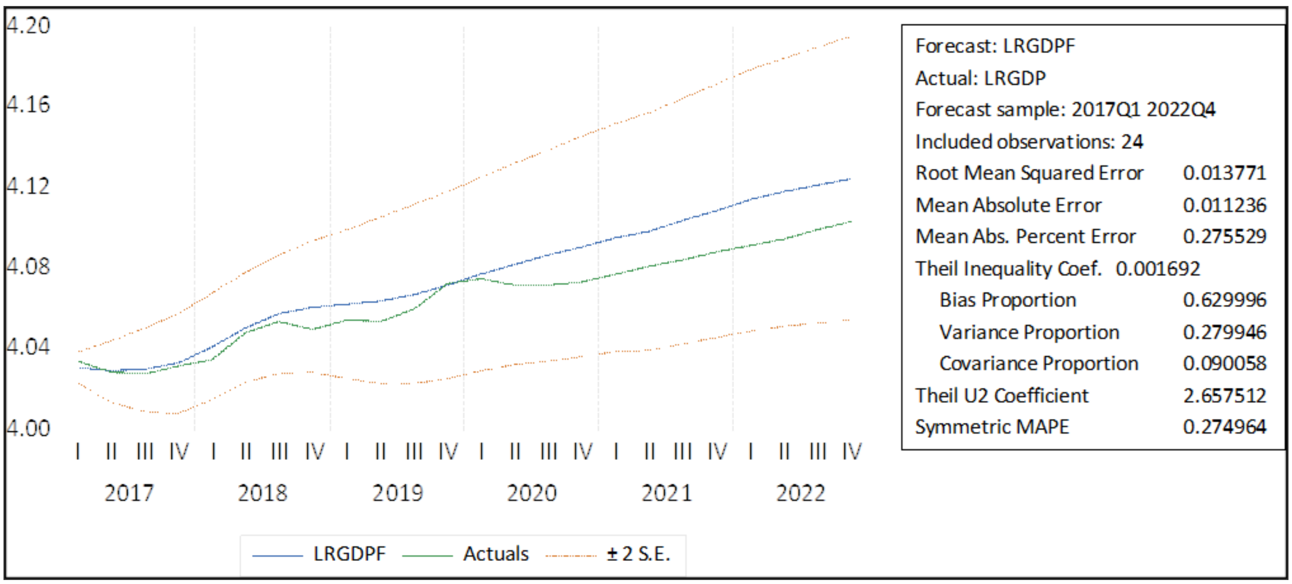

4.8
4.4
4.0
3.6
3.2
2.8
I II III IV I II III IV I II III IV I II III IV I II III IV I II III IV
2017 2018 2019 2020 2021 2022
LTGDPF
± 2 S.E.
Forecast: LTGDPF
Actual: LTGDP
Forecast sample: 2017Q1 2022Q4
Included observations: 24
Root Mean Squared Error 0.119951
Mean Absolute Error 0.099489
Mean Abs. Percent Error 2.725583
Theil Inequality Coef. 0.015479
Bias Proportion 0.493400
Variance Proportion 0.128147
Covariance Proportion 0.378453
Theil U2 Coefficient 0.393651
Symmetric MAPE 2.669249

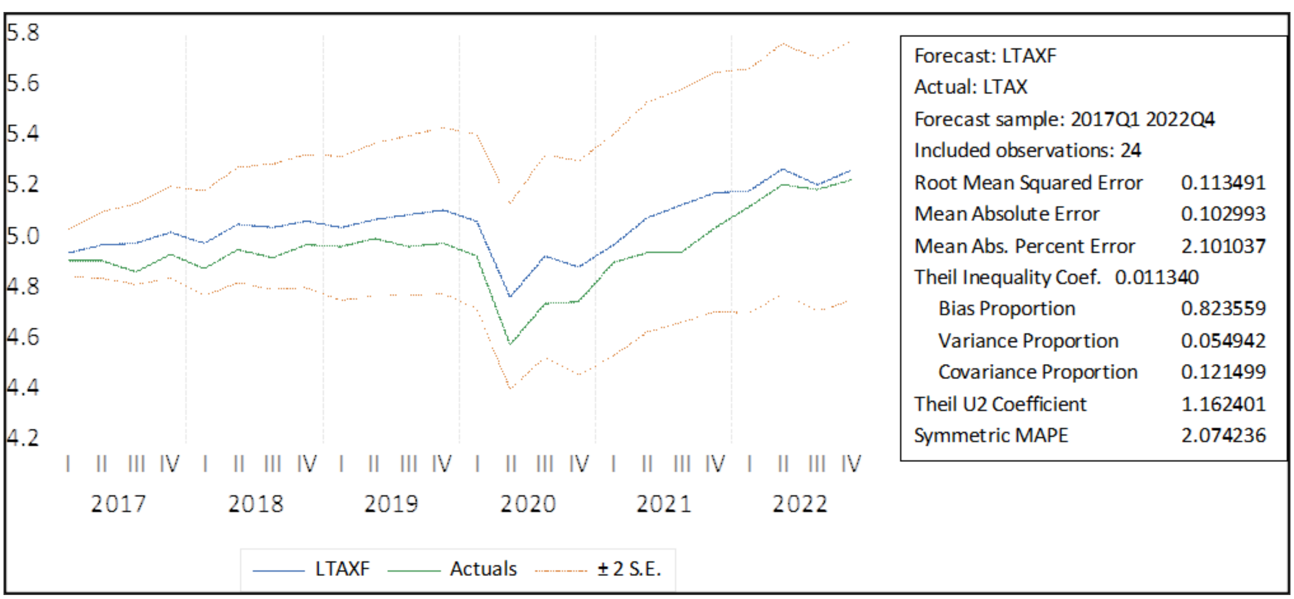

Forecast: LPGDPF
Actual: LPGDP
Forecast sample: 2017Q1 2022Q4
Included observations: 24
Root Mean Squared Error 0.156839
Mean Absolute Error 0.129309
Mean Abs. Percent Error 3.332531
Theil Inequality Coef. 0.020399
Bias Proportion 0.298257
Variance Proportion 0.055186
Covariance Proportion 0.646558
Theil U2 Coefficient 0.769603
Symmetric MAPE 3.389956
4.8
4.4
4.0
3.6
3.2
2.8
I II III IV I II III IV I II III IV I II III IV I II III IV I II III IV
2017 2018 2019 2020 2021 2022
LPGDPF Actuals ± 2 S.E.

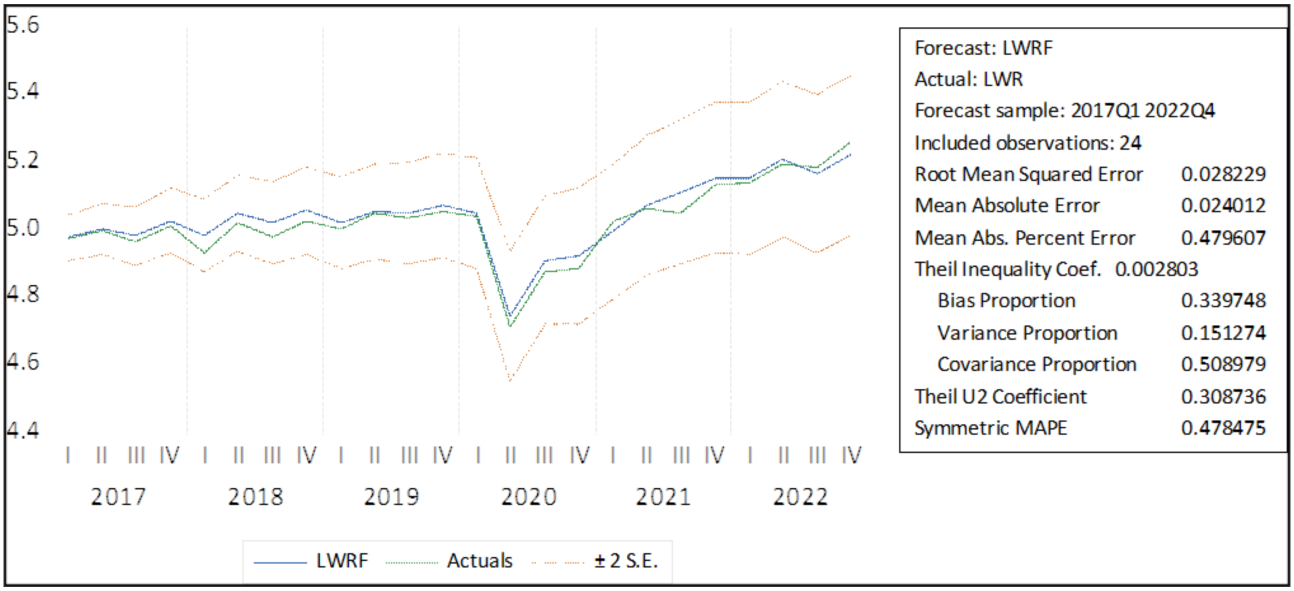

Forecast: LWGDPF
Actual: LWGDP
Forecast sample: 2017Q1 2022Q4
Included observations: 24
Root Mean Squared Error 0.057044
Mean Absolute Error 0.043465
Mean Abs. Percent Error 2.020641
Theil Inequality Coef. 0.013240
Bias Proportion 0.072340
Variance Proportion 0.003571
Covariance Proportion 0.924089
Theil U2 Coefficient 1.066882
Symmetric MAPE 2.006713
2.8
2.6
2.4
2.2
2.0
1.8
1.6
I II III IV I II III IV I II III IV I II III IV I II III IV I II III IV
2017 2018 2019 2020 2021 2022
LWGDPF Actuals ± 2 S.E.


Forecast: LOPSF
Actual: LOPS
Forecast sample: 2017Q1 2022Q4
Included observations: 24
Root Mean Squared Error 0.092852
Mean Absolute Error 0.081327
Mean Abs. Percent Error 1.768451
Theil Inequality Coef. 0.009995
Bias Proportion 0.750998
Variance Proportion 0.021183
Covariance Proportion 0.227819
Theil U2 Coefficient 3.471541
Symmetric MAPE 1.748301
5.6
5.4
5.2
5.0
4.8
4.6
4.4
4.2
4.0
I II III IV I II III IV I II III IV I II III IV I II III IV I II III IV
2017 2018 2019 2020 2021 2022
LOPSF Actuals ± 2 S.E.


Forecast: LEWF
Actual: LEW
Forecast sample: 2017Q1 2022Q4
Included observations: 24
Root Mean Squared Error 0.198476
Mean Absolute Error 0.169624
Mean Abs. Percent Error 5.879197
Theil Inequality Coef. 0.033257
Bias Proportion 0.107938
Variance Proportion 0.002274
Covariance Proportion 0.889788
Theil U2 Coefficient 0.684178
Symmetric MAPE 5.733828
2017 2018 2019 2020 2021 2022
LEWF Actuals ± 2 S.E.


Forecast: LRGDPF
Actual: LRGDP
Forecast sample: 2017Q1 2022Q4
Included observations: 24
Root Mean Squared Error 0.013771
Mean Absolute Error 0.011236
Mean Abs. Percent Error 0.275529
Theil Inequality Coef. 0.001692
Bias Proportion 0.629996
Variance Proportion 0.279946
Covariance Proportion 0.090058
Theil U2 Coefficient 2.657512
Symmetric MAPE 0.274964
2017 2018 2019 2020 2021 2022
LRGDPF Actuals ± 2 S.E.


Forecast: LTAXF
Actual: LTAX
Forecast sample: 2017Q1 2022Q4
Included observations: 24
Root Mean Squared Error 0.113491
Mean Absolute Error 0.102993
Mean Abs. Percent Error 2.101037
Theil Inequality Coef. 0.011340
Bias Proportion 0.823559
Variance Proportion 0.054942
Covariance Proportion 0.121499
Theil U2 Coefficient 1.162401
Symmetric MAPE 2.074236
2017 2018 2019 2020 2021 2022
LTAXF Actuals ± 2 S.E.


5.6
5.4
5.2
5.0
4.8
4.6
4.4
I II III IV I II III IV I II III IV I II III IV I II III IV I II III IV
2017 2018 2019 2020 2021 2022
LWRF
Actuals
± 2 S.E.
Forecast: LWRF
Actual: LWR
Forecast sample: 2017Q1 2022Q4
Included observations: 24
Root Mean Squared Error 0.028229
Mean Absolute Error 0.024012
Mean Abs. Percent Error 0.479607
Theil Inequality Coef. 0.002803
Bias Proportion 0.339748
Variance Proportion 0.151274
Covariance Proportion 0.508979
Theil U2 Coefficient 0.308736
Symmetric MAPE 0.478475