International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(8s) 2026 pp. 964-979
https://doi.org/10.70917/ijcisim-2026-3377
Article

Analysis of the factors affecting the general budget deficit in Iraq using stepwise regression, hierarchical regression, and artificial intelligence models

Shreen Ali Hussein[1], Saad Kadhim Hamza[2]

[1]Department of Quality Assurance and University Performance, University of Baghdad, Iraq
[2]Department of Statistics, College of Administration and Economics, University of Baghdad, Iraq
Emails:shreen.a@uobaghdad.edu.iq , Saad.hamza@coadec.uobaghdad.edu.iq

Abstract: The public budget deficit is a recurring structural problem in rentier economies, but the factors affecting the public budget deficit have yet to be quantitatively confirmed in the application of literature in the case of Iraq. The aim of this study is to reveal the most influential factors affecting the Iraqi public budget deficit and to order them in descending order based on their impact based on an analytical framework that combines traditional statistical analysis with machine learning models, and monthly data with the number of observations equal to 194 and the number of economic and financial data obtained from official sources, including the Central Bank of Iraq, the Ministry of Finance, and the Ministry of Planning. The research was carried out in two phases. The first step was to conduct diagnostic tests to determine whether the linear regression assumptions held true; the results suggested that there was no significant issues of multicollinearity as all the Variance Inflation Factor (VIF) values were below the critical threshold. The second stage comprised the estimation of three linear models, including the full model, stepwise regression model based on the Akaike Information Criterion (AIC), and hierarchical regression model with an 80/20 training–testing data split, and the comparison of these models with two machine learning models (Random Forest and XGBoost). The results indicated that stepwise regression model was significant, with a smaller number of variables than the full model, but also showed superior out of sample predictive accuracy as compared to the other models. In addition, the relative importance of four variables was strongly agreed by the models, with oil revenue, actual spending, domestic investment and non-oil revenue being most important in explaining the differences among the public budgets. Others, however, such as public debt, showed some variations in the importance of the variable between the linear and the nonlinear models, indicating that there are more complex relationships that could be investigated. The study finds that combining traditional regression methods with measures of variable importance from machine learning algorithms gives a more powerful tool for determining the factors behind the budget deficit which increases the reliability of the results when cross-validating the findings in various ways from different analytical methods.

Keywords: Budget Deficit, Stepwise Regression, Hierarchical Regression, Random Forest, XGBoost, Variable Selection, Fiscal Policy, Model Consensus.

1. Introduction

One of the indices that indicates the effectiveness of fiscal policy and a government's ability to balance its financial resources with financial expenditures is the public budget deficit. In the economic literature this has been a subject of much discussion because of its relevance for macroeconomic stability both directly and indirectly. Besides affecting investment, economic growth and monetary stability in the long run, high budget deficits can result in a greater use of public borrowing and growth of government debt [16,17]. In addition, budget deficits are one of the largest issues of the highly dependent developing and rentier economies relying on unpredictable and unstable revenue streams.The Iraqi economy is a good example of this kind of economy because oil revenues are the main source of financing for the budget of the state. In recent years, Iraq's public finances have been under mounting strain, due to

the frequent volatility of oil prices in the international market and the ever-increasing burden of government spending, as well as the rising costs of economic and service activities of the government. All of these factors have directly influenced the size and trend of the fiscal deficit and therefore the study of the determinants of the fiscal deficit and the analysis of the factors affecting it is scientific and practical necessity, in order to understand the nature of the fiscal imbalances of the Iraqi economy. The literature of economics suggests that the budget deficit is not a single variable but rather is affected by a network of economic, financial and monetary variables. Some of the most significant of these variables are public debt, government spending, oil and non-oil revenues, investment, exchange rate, foreign exchange reserves, excess reserves, and monetary and financial sector variables. The impact of these factors on the budget deficit depends on the structural features of the economy and the time periods considered, however. Thus, the results of the previous studies cannot be used directly in analyzing the Iraqi case. In past studies, the determinants of fiscal deficits have been studied mainly using traditional regression models, which have helped to improve the understanding of the relationship between different economic variables. In many instances, however, the linearity assumption implicit in these models may not be sufficient, especially in situations where the economic relationships are complex and/or nonlinear effects are present that cannot be adequately represented by the traditional models. On the contrary, machine learning techniques have shown significant potential in economic and financial research in recent years, as they are able to identify complex patterns, manage nonlinear relationships and extract information from data efficiently [2,3,6,7]. While a growing number of studies have used traditional regression-based methods or machine learning techniques, there is a few studies that combine both techniques in a single analytical framework to obtain the determinants of the public budget deficit in Iraq. Hence, the research gap that is focused on in this study is the integration of a group of statistical methods and artificial intelligence techniques and the comparison of their results in a single holistic framework. Therefore, this study seeks to study the factors that impact the public budget deficit in Iraq for the period January 2008 to January 2024, on a monthly basis. This is done using some of the most widely used variable selection methods in linear models such as stepwise regression and hierarchical regression, and among the most widely used machine learning models in economic data analysis are the Random Forest and XGBoost algorithms. The study also aims to determine which variables are most important in causing the budget deficit and to compare the results of the separate models, as well as to study how close the models are in their rankings of the variables and in their explanations of the budget deficit.

The significance of this study lies in the fact that it goes beyond merely estimating the effects of economic and financial variables. Rather, it presents an analytical framework that integrates traditional statistical inference with modern machine learning techniques, thereby providing a more comprehensive understanding of the determinants of the public budget deficit in Iraq and contributing to the support of decisions related to fiscal policy and public resource management.

1.1 Research Problem

The public budget deficit is an economic problem, which is affected by a set of interrelated financial, monetary and economic variables. Many studies have investigated the causes of fiscal deficits, but the results have been inconsistent on what variables are the most important and significant. This mismatch is partly due to the different analytical methods used and because most studies have been based on the use of traditional statistical models, which have not been able to fully exploit recent developments in the field of artificial intelligence and machine learning. For Iraq, empirical evidence on the ranking of factors affecting the budget deficit is still limited and the comparison of traditional regression models with artificial intelligence models is still relatively limited in this context. Thus, the research problem is focused on answering the following question: What are the most influential factors affecting the public budget deficit in Iraq, and what is the difference in the determinants identified based on the method used (stepwise regression, hierarchical regression and models of artificial intelligence)?

1.2 Research Objectives

The purpose of this study is to analyze the factors affecting the public budget deficit in Iraq using statistical methods and modern techniques of artificial intelligence to examine the nature of the relationship between the fiscal deficit and public debt, the actual expenditure, money supply, foreign exchange reserves, investment, excess reserves, the exchange rate and public revenue. The study also aims to compare the capacity of stepwise regression, hierarchical regression and the models of Random Forest and XGBoost to determine the variables that have the most impact on the budget deficit and to rank them by their relative impact. In addition, it seeks to compare the performance of these models in explaining the phenomenon under study and eventually to establish a framework of analysis to provide a more accurate understanding of the determinants of the fiscal deficit in Iraq and to assist in financial and economic decision making.

1.3 Research Importance

The importance of this research stems from the significance of the public budget deficit as one of the key indicators of financial and economic stability. The study is also important because it employs a variety of analytical approaches that allow the phenomenon to be examined from multiple perspectives rather than relying on a single model. From an applied perspective, the findings provide quantitative evidence that can assist policymakers in identifying the variables most closely associated with the fiscal deficit, thereby supporting the formulation of more effective fiscal and economic policies.

1.4 Research Contribution

The scientific contribution of this study lies in integrating traditional statistical regression methods with artificial intelligence models within a single analytical framework to investigate the determinants of the public budget deficit in Iraq. The study also contributes by providing a direct comparison of the ability of these models to identify influential variables and rank them according to their relative importance. The findings are expected to offer a deeper understanding of the factors governing the budget deficit, while also providing empirical evidence on the potential of artificial intelligence models in addressing economic and financial issues.

1.5 Literature Review

The public budget deficit is one of the issues that has attracted considerable attention from researchers in macroeconomics and public finance due to its implications for financial and monetary stability as well as economic growth. The literature indicates that the persistence of high budget deficits increases the need for government borrowing and leads to the accumulation of public debt, in addition to its repercussions on investment, savings, and overall economic activity [16,17]. Consequently, numerous studies have sought to identify the factors responsible for the emergence and escalation of fiscal deficits and to uncover the nature of the relationships between economic and financial variables and public budget performance.

The findings of previous studies indicate that the determinants of the budget deficit are not confined to a single aspect of economic activity; rather, they encompass a broad range of financial, monetary, and economic variables. Empirical reviews have shown that public debt, economic growth, inflation, trade openness, government expenditure, and public revenues are among the most prominent factors associated with fiscal deficits. However, the direction and magnitude of their effects vary according to the structure of the economy and the prevailing institutional conditions [17]. Other studies have also demonstrated that an expanding government sector and rising public debt servicing costs are often associated with higher levels of fiscal deficits, whereas economic growth and improvements in public revenues contribute to alleviating the pressures on the public budget [16]. In developing economies, monetary and financial variables assume greater importance due to the limited availability of public revenue sources and governments' reliance on borrowing or indirect financing to cover budget deficits. In this context, the literature has shown that money supply, exchange rates, foreign exchange reserves, and investment levels can directly or indirectly affect fiscal deficits through their influence on government revenues, expenditure levels, or overall economic activity. Empirical evidence also suggests that the relationships among these variables may not remain stable over time, which explains why study findings often differ across countries depending on the characteristics of the economy under investigation.

In Iraq, the study of budget deficit determinants is of particular importance due to the rentier nature of the economy and its heavy dependence on oil revenues as the primary source of financing public expenditure. Numerous studies have demonstrated that fluctuations in oil revenues and global oil prices are among the most significant factors affecting public budget performance, while other variables such as government expenditures, public debt, exchange rates, investment, and foreign exchange reserves interact to varying degrees in explaining changes in the fiscal deficit [16,17]. The literature has also emphasized that the continued heavy reliance on oil revenues makes public finances more vulnerable to external shocks and global economic fluctuations, thereby increasing the importance of identifying the most influential determinants of the fiscal deficit and assessing their relative roles. Despite the significant contributions of previous studies to understanding the determinants of budget deficits, most of them have relied on traditional regression models or focused on a limited set of economic variables. In contrast, recent years have witnessed a remarkable expansion in the use of machine learning techniques in economic and financial research due to their ability to handle nonlinear relationships and complex interactions among variables [2,3,6,7]. Therefore, integrating traditional statistical regression methods with modern machine learning techniques provides a more comprehensive framework for analyzing the determinants of the budget deficit and identifying the variables that exert the greatest influence on it.

1.6 Statistical and Artificial Intelligence Approaches

Most previous studies have relied on traditional econometric models to analyze the determinants of budget deficits, employing linear regression models, time-series analysis, and cross-sectional and panel data models to estimate the effects of economic and financial variables on fiscal deficits. These models have contributed significantly to explaining important aspects of economic relationships; however, their effectiveness remains dependent on a set of statistical assumptions that may limit their performance in the presence of nonlinear relationships or complex interactions among variables.

With the rapid advancement of machine learning applications, many recent studies have turned to more flexible models for economic and financial analysis. Among the most widely used models in this field are Random Forest and XGBoost, owing to their ability to handle nonlinear relationships and to extract the relative importance of variables with a high degree of accuracy. Empirical studies in the areas of economic forecasting, financial markets, and risk management have demonstrated that these models often achieve superior predictive performance compared with traditional models. Moreover, they possess the capability to uncover complex patterns that may be difficult to detect using conventional regression techniques.

1.7 Research Gap

Although numerous studies have examined the determinants of public budget deficits, most have focused on applying a single analytical approach, whether traditional econometric models or modern machine learning techniques. Furthermore, studies that directly compare variable-selection methods in conventional regression models with artificial intelligence models remain limited, particularly in developing economies in general and the Iraqi economy in particular. Therefore, the research gap lies in the absence of a comprehensive study that integrates stepwise regression, hierarchical regression, Random Forest, and XGBoost within a unified comparative framework aimed at identifying and ranking the factors affecting the public budget deficit in Iraq. The present study seeks to address this gap by evaluating the ability of these models to explain the fiscal deficit and to identify the most influential variables affecting it.

2. Data and Methodology

2.1 Data Description

The study utilized monthly data for the Iraqi economy obtained from statistical bulletins and official reports issued by the Central Bank of Iraq. The dataset covers the period from January 2008 to January 2024 and consists of 194 monthly observations. This period was selected based on the availability of consistent monthly data for all variables under investigation, thereby ensuring the coherence and reliability of the database used in the analysis.

The public budget deficit (Y) represents the dependent variable in this study. The explanatory variables include total public debt (X1), actual expenditures (X2), currency reserves (X3), investment (X4), excess reserve (X5), the exchange rate of the U.S. dollar against the Iraqi dinar (X6), oil revenues (X7), foreign reserves (X8), other revenues (X9), and foreign investments (X10). These variables were selected based on the economic literature related to the determinants of public budget deficits. In addition, the opinions of several specialists in economics and public finance were considered to ensure that the selected variables are appropriate for the nature and characteristics of the Iraqi economy.

The analysis was limited to variables for which regular monthly time series were available throughout the study period. Several other economic variables discussed in the previous literature were excluded due to the lack of consistent monthly data. This procedure was adopted to maintain the homogeneity of the dataset and to avoid methodological issues associated with combining data of different frequencies. Consequently, it enhances the reliability of the findings derived from both the statistical models and the artificial intelligence models employed in the study.

2.2 Selection of the Best Regression Equation

The selection of explanatory variables to be included in a regression model is considered one of the most challenging issues in regression analysis. In practice, not all variables possess the same level of importance; a model may contain explanatory variables that are correlated with one another, and the omission of important variables or the inclusion of less relevant ones may reduce the accuracy of the estimates and weaken the overall statistical significance of the model. For this reason, researchers have proposed a variety of statistical techniques aimed at identifying the

best subset of explanatory variables, thereby achieving a balance between the goodness of fit of the model and its parsimony in terms of the number of parameters. Among the most prominent of these techniques are:

2.2.1 Stepwise Regression

Stepwise Regression is one of the most widely used variable-selection techniques in multiple regression models. It aims to construct an efficient model by identifying the variables that are most capable of explaining variations in the dependent variable while excluding variables with limited contributions [11,13]. This approach relies on an iterative procedure that combines the inclusion and removal of explanatory variables according to specific statistical criteria, ultimately resulting in the model that best fits the data [8]. The algorithm begins by examining the relationship between the dependent variable Y and each explanatory variable individually. The variable that exhibits the greatest explanatory power or the strongest association with the dependent variable is then entered into the model. After the first variable is included, the model is re-estimated and the remaining variables are evaluated to identify the one that contributes most to improving the model. At each stage, the statistical significance of the variables already included in the model is reassessed. If a previously selected variable loses its statistical significance as a result of adding a new variable, it is removed from the model [5,11]. The process of adding and removing variables continues iteratively based on tests of statistical significance and model goodness-of-fit criteria until a final model is obtained that cannot be further improved by either adding a new variable or removing any of the existing variables. This approach is distinguished by its ability to reduce the number of explanatory variables while retaining those with the greatest influence, thereby contributing to the development of a more parsimonious and interpretable model [13].

The study adopted the following multiple linear regression model as the basis for the variable selection procedures:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \cdots + \beta_{10} X_{10} + \varepsilon \ldots\ldots (1)$$

Where Y represents the public budget deficit, X1–X10 represent the explanatory variables, and ε denotes the random error term. The application of stepwise regression in this study aims to identify the variables that exert the greatest influence on the Iraqi public budget deficit and to derive the model that achieves the best balance between explanatory power and model parsimony. The following flowchart illustrates the procedure of the method [9,12,18].

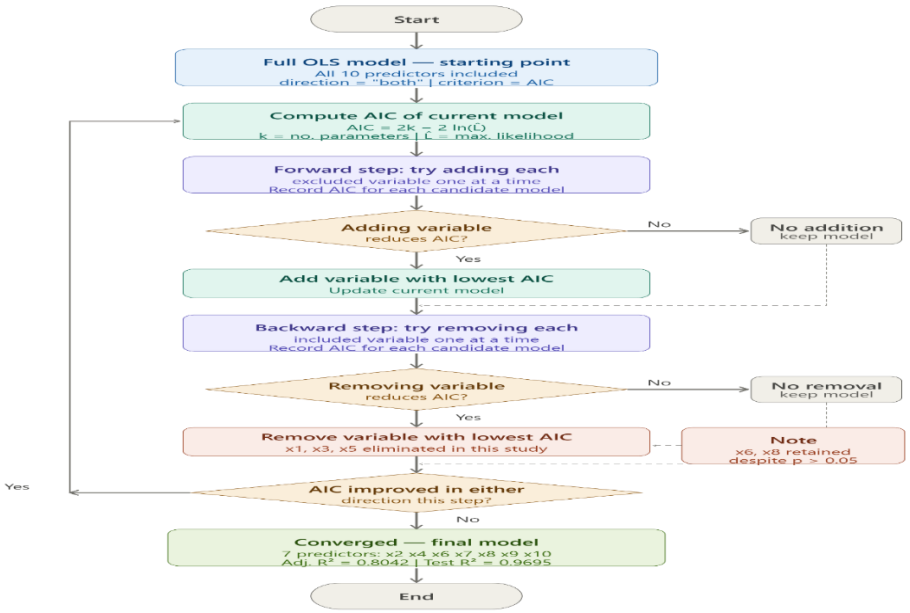

Figure (1). Steps for Building the Model Using the Stepwise Regression Method.

2.2.2 Hierarchical Regression

Hierarchical Regression is one of the methods used to determine the importance of explanatory variables affecting the dependent variable. It is often regarded as a methodological alternative to stepwise regression when the researcher has a clear theoretical basis for ordering the variables included in the model [15,14]. This approach involves

entering explanatory variables sequentially in the form of steps or blocks; however, it differs from stepwise regression in that the order of variable entry is not determined by an automatic statistical algorithm but rather by scientific theory, previous studies, and the researcher's domain expertise [10,14]. Kerlinger noted that there is no single ideal method for determining the order in which variables should be entered. Instead, a thorough understanding of the research problem and the theoretical framework underlying the study provides the most logical basis for specifying the sequence of variable entry into the regression model [10]. Similarly, Fox argued that although automated model-selection procedures may sometimes compensate for certain shortcomings in the data, they cannot substitute for the researcher's theoretical knowledge and scientific judgment in selecting the most appropriate variables [5]. Therefore, hierarchical regression is considered an appropriate tool for analyzing explanatory relationships when the objective is to examine the additional contribution of each set of variables in explaining the variation in the dependent variable [4,15]. This approach is based on comparing successive models by examining changes in the coefficient of determination or the adjusted coefficient of determination after the inclusion of each variable or group of variables. This enables researchers to evaluate the extent of improvement contributed by each step to the model [14,19]. One of the important features of hierarchical regression is that it combines the computational power of statistical software with the scientific expertise of the researcher. The grouping of variables and the order of their entry are determined based on theoretical foundations and accumulated findings from previous studies, making it more closely aligned with the scientific logic of the study than automatic variable-selection methods [15,10]. For this reason, hierarchical regression is widely used in applied research aimed at explaining economic and social phenomena and determining the relative importance of different groups of variables [1,4,19].

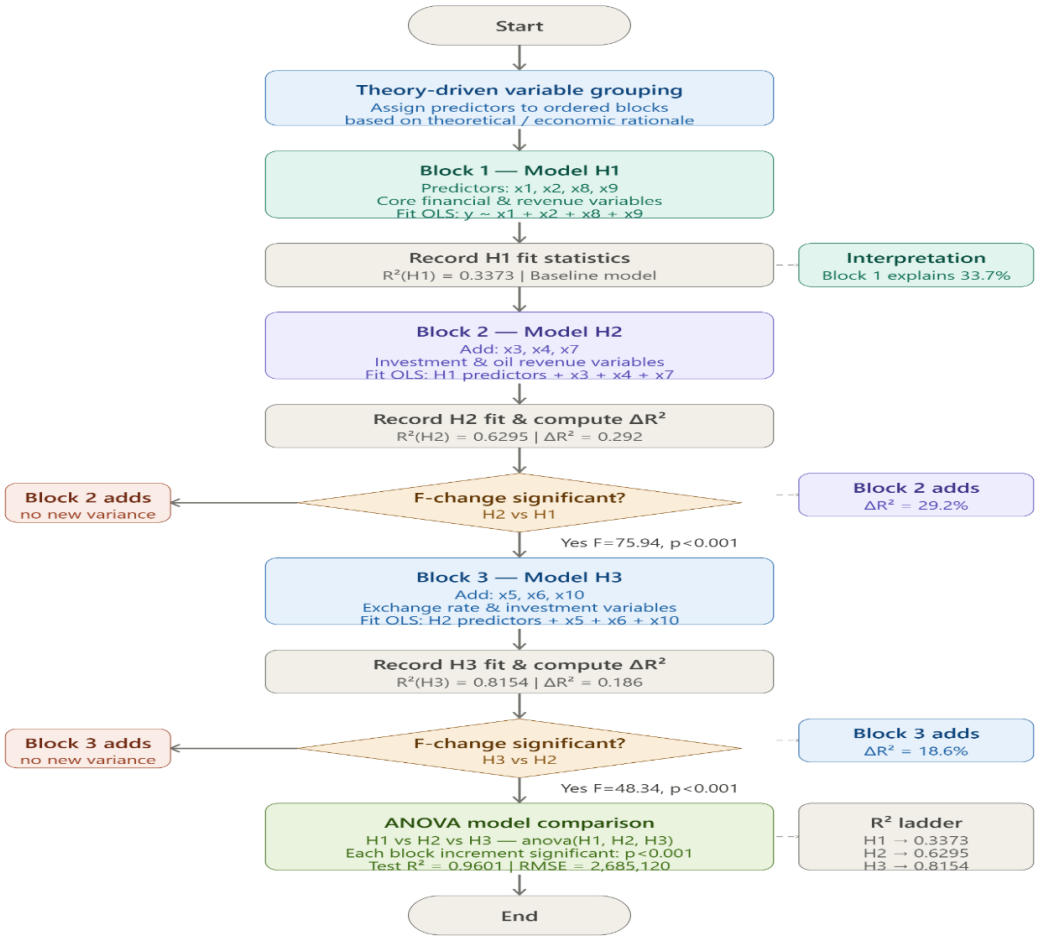

Figure (2). Steps for Building the Model Using the Hierarchical Regression Method.

2.2.3 Random Forest

Random Forest is one of the most widely used machine learning algorithms based on the Ensemble Learning approach. It was proposed by Leo Breiman in 2001 with the objective of improving predictive accuracy and reducing the high variance problem associated with individual decision trees [2]. The algorithm operates by constructing a large number of decision trees using random samples drawn from the original dataset through the bootstrap sampling technique, while a random subset of explanatory variables is selected at each splitting node within the tree [2,7]. The fundamental idea behind the model is to combine the predictions of a large number of trees rather than relying on a single tree. In regression problems, the predicted value is obtained by averaging the predictions generated by all trees. This procedure helps reduce variance and enhances the predictive performance of the model compared with traditional approaches [2]. One of the key advantages of Random Forest is its ability to handle nonlinear relationships and complex interactions among economic and financial variables without requiring the specification of a particular functional form. In addition, it provides quantitative measures of Variable Importance, enabling researchers to identify the variables that have the greatest influence on the dependent variable [7].

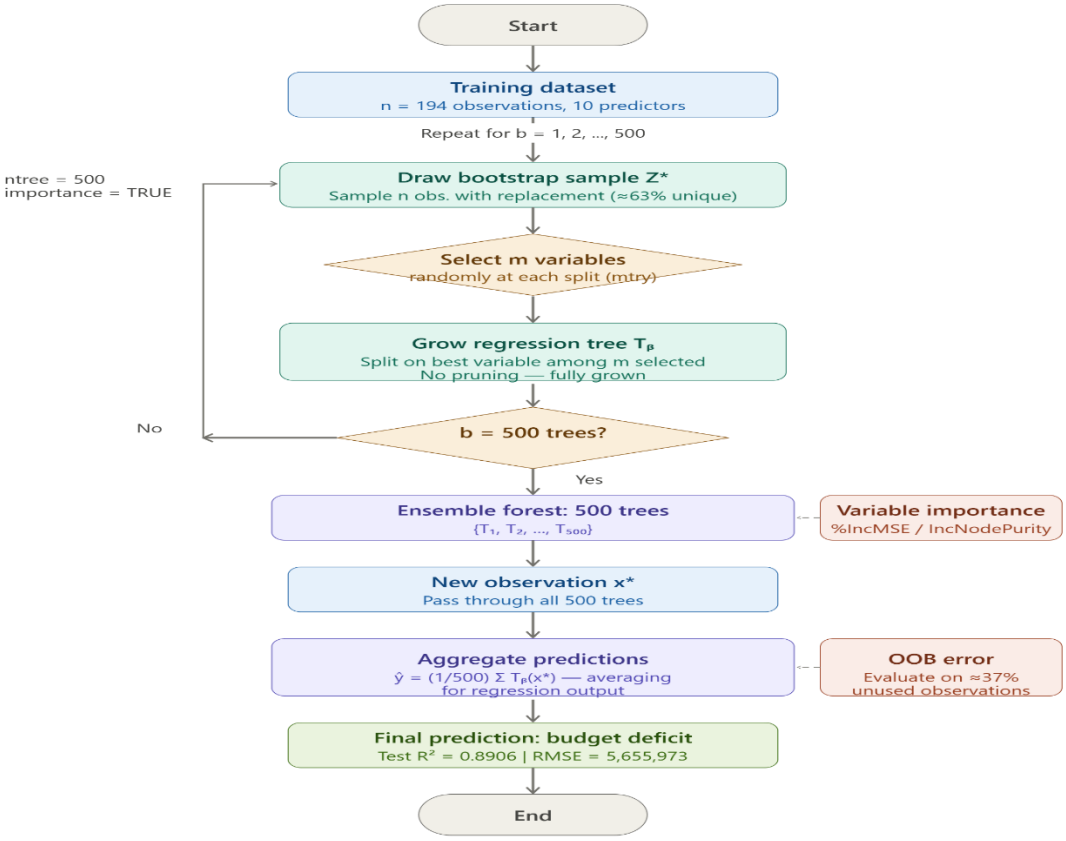

Figure (3). Steps for Building the Model Using the Random Forest Method.

2.2.4 XGBoost Model

The XGBoost (Extreme Gradient Boosting) model is considered one of the most efficient machine learning algorithms for prediction and classification tasks. It was developed by Tianqi Chen and Carlos Guestrin with the aim of improving the efficiency of traditional gradient boosting algorithms by combining high computational speed with strong predictive performance [3]. The model is based on the Gradient Boosting framework introduced by Jerome H. Friedman, which constructs a sequence of decision trees in which each new model attempts to correct the residual errors of the previous model, thereby progressively improving prediction accuracy [6].

The objective function in the XGBoost model is expressed as follows:

$$L = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{k} \Omega(f_k) \ldots\ldots (2)$$

Where:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \|w\|^2 \ldots\ldots (3)$$

The first term represents the loss function, which measures the discrepancy between the actual values and the predicted values. The second term represents the regularization term, which aims to control the complexity of the model and reduce overfitting by imposing a penalty on overly complex decision trees [3]. Regularization is considered one of the most important features that distinguishes XGBoost from many other boosting algorithms, as it enhances the model's ability to generalize when applied to new data that were not used during the training process [3,7]. In addition, the model utilizes both the first-order and second-order derivatives of the loss function when updating its parameters, which increases the efficiency of the optimization process and accelerates convergence toward the optimal solution compared with traditional boosting algorithms [3].

XGBoost is also distinguished by its ability to capture nonlinear relationships and complex interactions among variables. In addition, it supports parallel processing and the automatic handling of missing values, making it well suited for the analysis of large-scale and high-dimensional datasets [3,7]. The model also provides several measures of variable importance, such as Gain, Cover, and Frequency, which help identify the variables that exert the greatest influence on the dependent variable and facilitate a clearer interpretation of the model's results [3]. In this study, the XGBoost model was employed to identify the most important factors affecting the Iraqi public budget deficit. The model's hyperparameters, particularly the learning rate, tree depth, and regularization parameters, were optimized using a cross-validation approach in order to achieve the best possible predictive performance [3].

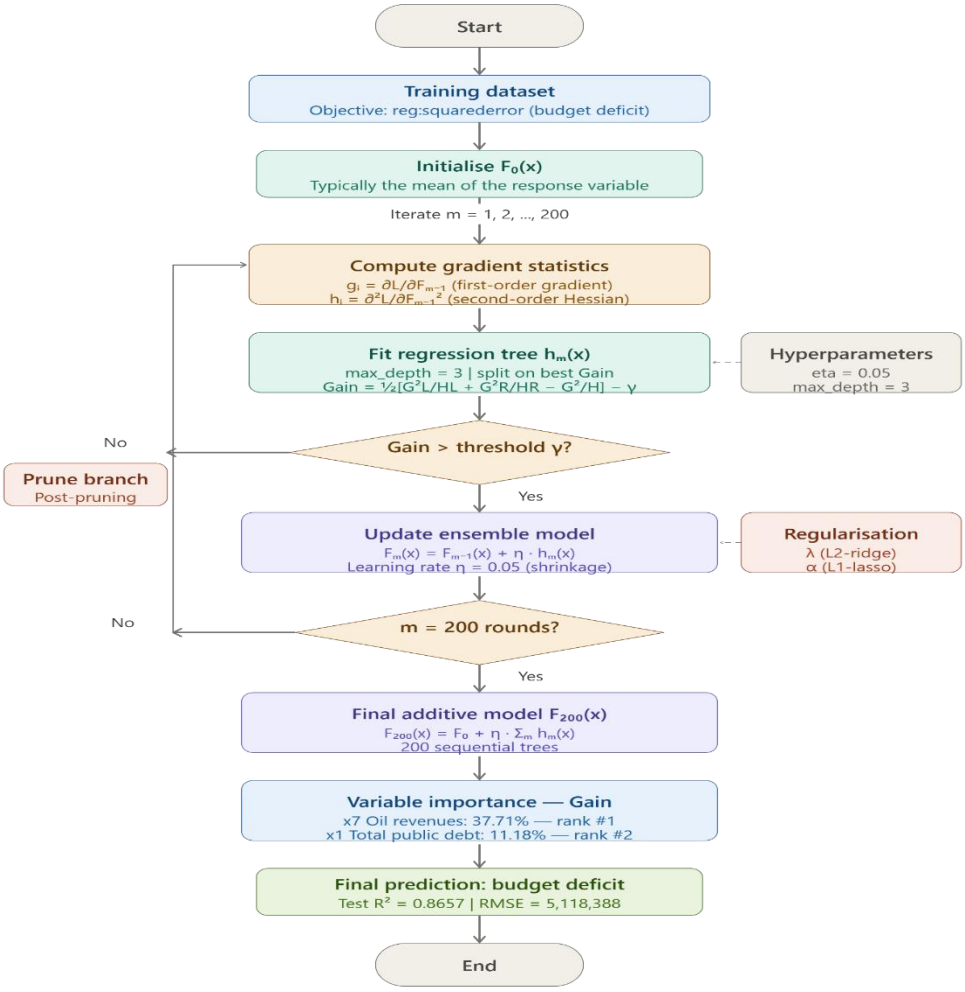

Figure (4) illustrates the steps involved in building the model using the XGBoost approach.

2.3 Study Variables

The study consists of one dependent variable, namely the Iraqi budget deficit (Y), and ten independent variables representing the major fiscal, monetary, and investment determinants, as presented in Table (1).

Table (1). Definition of the Study Variables

Symbol	Description	Role
y	**Budget Deficit**	**dependent variable**
x1	**Total Public Debt**	**independent variables**
x2	**Actual Expenses**	**independent variables**
x3	**Currency Reserves**	**independent variables**
x4	**Investment**	**independent variables**
x5	**Excess Reserve**	**independent variables**

x6	**Exchange Rate**	**independent variables**
x7	**Oil Revenues**	**independent variables**
x8	**Foreign Reserves**	**independent variables**
x9	**Other Revenues**	**independent variables**
x10	**Foreign Investments**	**independent variables**

2.4 Diagnostic Tests for the Regression Model

Prior to model estimation, diagnostic tests were conducted to verify the assumption of the absence of severe multicollinearity among the explanatory variables. After estimation, additional diagnostic tests were performed to examine the absence of serial correlation in the residuals and heteroscedasticity, as these are among the most important problems that may affect the validity and reliability of the estimated results.

First: Variance Inflation Factor (VIF) Test for Multicollinearity.

Multicollinearity was assessed using the Variance Inflation Factor (VIF) and the Tolerance coefficient. According to the criteria commonly adopted in the statistical literature, a VIF value exceeding 10 indicates severe multicollinearity that may compromise the validity of the model estimates. Values falling between 5 and 10 are considered a cautionary zone warranting attention, while values below 5 indicate that multicollinearity is not a substantive concern.

Table (2). Multicollinearity Statistics for the Model Variables

Variable	Tolerance	VIF	Sig.	Decision
Total Public Debt	**0.453**	**2.205**	**< 0.001**	**Acceptable**
Actual Expenditures	**0.257**	**6.370**	**< 0.001**	Caution
Foreign Exchange Reserves	**0.256**	**3.908**	**0.710**	**Acceptable**
Surplus Reserve	**0.423**	**2.363**	**0.044**	**Acceptable**
Exchange Rate	**0.432**	**2.315**	**0.700**	**Acceptable**
Oil Revenues	**0.279**	**5.579**	**< 0.001**	Caution
Foreign Reserve	**0.279**	**3.579**	**0.004**	**Acceptable**
Other Revenues	**0.502**	**1.992**	**0.658**	**Acceptable**
Foreign Investments	**0.350**	**2.855**	**0.207**	**Acceptable**
Investment	**0.415**	**2.752**	**0.635**	**Acceptable**

The results presented in Table (2) indicate that all Variance Inflation Factor (VIF) values are below the critical threshold of 10, confirming the absence of severe multicollinearity. Two variables exceeded the cautionary threshold of 5: actual expenditures (VIF = **6.370**) and oil revenues (VIF = **5.579**). Nevertheless, both values remain well below the critical threshold of 10, and their retention is justified by their strong theoretical relevance and consistently high statistical significance ($p < 0.001$) across all model specifications **.**

Second: Autocorrelation Test (Breusch–Godfrey Test)

To verify the assumption of error independence and the absence of autocorrelation in the model residuals, the Breusch–Godfrey test was applied, as it is considered one of the most appropriate tests for multiple regression models. The results showed that the p-values of the test at the first and second lag orders were 0.301 and 0.379, respectively,

both of which are greater than the significance level of 0.05. This indicates that there is no statistically significant evidence of autocorrelation in the residuals. Accordingly, the null hypothesis of independent random errors cannot be rejected, supporting the validity of the model estimates and enhancing the reliability of the statistical inferences derived from them.

Third: Homoscedasticity Test of Random Errors (Breusch–Pagan Test)

To verify the assumption of homoscedasticity of the random errors, the Breusch–Pagan test was employed to detect the presence of heteroscedasticity. The results indicated that the associated p-value was 0.327, which exceeds the adopted significance level of 0.05, suggesting that there is no statistical evidence of heteroscedasticity in the residuals. This finding implies that the variance of the error terms remains stable across different observations, thereby satisfying one of the key assumptions of the Ordinary Least Squares (OLS) method and enhancing the efficiency and reliability of the statistical estimates obtained from the model.

Fourth: Unit Root Test (Augmented Dickey–Fuller Test, ADF)

Given that the study relies on monthly time-series data, the Augmented Dickey–Fuller (ADF) test was conducted to examine the stationarity of the time series and to avoid the potential problem of spurious regression relationships. The results revealed that the dependent variable (budget deficit) is stationary at level, whereas several explanatory variables do not satisfy the stationarity condition at the same level. This finding is consistent with the economic nature of macroeconomic and financial variables, which often exhibit long-term trends. Since the primary objective of the study is to identify the most influential determinants of the budget deficit and to compare the predictive performance of different modelling approaches rather than to estimate long-run equilibrium relationships or cointegration vectors the variables were analysed in their original levels. This decision is supported by several considerations. First, the dependent variable (budget deficit) was found to be stationary at level, which is the primary condition for avoiding spurious regression in the context of a mixed-order system, as established by Pesaran et al. [20] in the bounds testing framework. Second, the Breusch–Godfrey test confirmed the absence of serial correlation in the model residuals ($p = 0.301$ and $p = 0.379$ at the first and second lag orders, respectively), and the Breusch–Pagan test confirmed homoscedasticity ($p = 0.327$). Together, these diagnostics indicate that the residuals satisfy the classical assumptions of the OLS estimator, which is a stronger practical guarantee against spurious inference than stationarity of the regressors alone. Third, the machine learning models (Random Forest and XGBoost) are entirely distribution-free and do not rely on stationarity assumptions, providing an additional layer of robustness to the conclusions drawn from the cross-model consensus analysis. Taken together, these considerations support the validity of the analytical approach adopted in this study and indicated the absence of econometric problems that could materially affect the study's results.

3. Results and Discussion

It should be noted that this chapter employs two types of coefficient of determination. The first is the Adjusted R^2, calculated using the training data (80% of the observations), which reflects the model's in-sample goodness of fit. The second is the Test R^2, calculated using the testing data (20% of the observations) according to R2=[cor(ytest,y^test)]2, which reflects the model's out-of-sample predictive performance. The evaluation of the models presented in the comparison table is based on Test R^2, as it is generally considered more appropriate for comparing models with different structures and estimation approaches. It should also be noted that all three linear models estimate their parameters using the Ordinary Least Squares (OLS) method; the difference among them lies in the strategy used for variable selection and the order in which variables are entered into the model, rather than in the estimation method itself.

3.1 Results of the Full Multiple Linear Regression Model

The application of the Multiple Linear Regression (Full Model) to the training dataset ($n = 155$) resulted in a model with a good level of fit, yielding an in-sample coefficient of determination of $R^2 = 0.8154$ and an Adjusted $R^2 = 0.8026$, indicating that the ten explanatory variables jointly explain approximately 80.3% of the variation in the budget deficit. The overall model was statistically significant, as the F-statistic exceeded its critical value and was significant at $p < 0.001$. The results of the t-tests identified five highly significant variables ($p < 0.001$): oil revenues (x7: $\beta = +0.952$), actual expenditures (x2: $\beta = -0.933$), investment (x4: $\beta = -7043$), foreign investments (x10: $\beta = +7044$), and other revenues (x9: $\beta = +1.959$). A notable finding is the identical absolute magnitude of the coefficients for x4 and x10, accompanied by opposite signs, suggesting a potential structural relationship between these variables that merits further investigation. In contrast, x1, x3, x5, x6, and x8 were not statistically significant at the conventional

significance level ($\alpha = 0.05$), which may be partly attributed to their overlap with stronger explanatory variables in the model. On the testing dataset (n = 39), the model achieved a Test R^2 of 0.9601, reflecting strong out-of-sample predictive performance; however, the prediction errors remained relatively high, with an RMSE of 2,685,120 and a MAPE of 176.00%.

3.2 Stepwise Regression Model

The bidirectional stepwise regression model was implemented in R using the step() function based on the Akaike Information Criterion (AIC), defined as $AIC = 2k - 2\ln(L)$, where k denotes the number of estimated parameters and ln(L) represents the log-likelihood function. At each iteration, the algorithm evaluates the addition or removal of variables in order to minimize the AIC value and terminates when no further improvement is possible. This criterion balances goodness of fit and model complexity and is generally regarded as more objective than relying solely on p-values for variable selection. The selection procedure resulted in a model containing seven explanatory variables (x2, x4, x6, x7, x8, x9, and x10), while x1, x3, and x5 were excluded. It is noteworthy that x6 ($p = 0.136$) and x8 ($p = 0.111$) were retained despite not achieving conventional statistical significance, reflecting the AIC philosophy of preserving variables that improve the overall trade-off between fit and parsimony even when their individual significance is relatively weak. In terms of in-sample performance, the model achieved $R^2 = 0.8131$ and Adjusted $R^2 = 0.8042$, slightly exceeding the corresponding value obtained from the full OLS model (Adjusted $R^2 = 0.8026$) despite utilizing only seven variables instead of ten, indicating that the removal of redundant variables improved the model's informational efficiency. The overall model was highly significant, with $F = 91.35$ ($df = 7, 147$; $p < 0.001$), exceeding the value obtained from the full regression model. When evaluated on the testing dataset, the model achieved a Test R^2 of 0.9695, an RMSE of 2,340,941, and a MAPE of 137.19%, representing the best predictive performance among all models considered in the study.

3.3 Hierarchical Multiple Regression Model.

The hierarchical regression approach was employed to examine the cumulative explanatory contribution of the variables through three sequential blocks constructed on theoretical and economic foundations.

Table (3): Changes in the Coefficient of Determination (R^2) Across the Hierarchical Regression Blocks.

Level	Variables Included	R^2	Adjusted R^2	ΔR^2	F-change	Sig.
H1 —First Block	**x1, x2, x8, x9**	**0.3373**	**—**	**—**	**—**	**—**
H2 – Second Block	**H1 + x3, x4, x7**	**0.6295**	**—**	**0.292**	**75.945**	**< 0.001**
H3 – Third Block	**H2 + x5, x6, x10**	**0.8154**	**0.8026**	**0.186**	**48.338**	**< 0.001**

In the first block (H1), which included x1, x2, x8, and x9, the model achieved an R^2 of 0.3373, indicating that these core variables explain only about one-third of the variation in the budget deficit. In the second block (H2), the addition of x3, x4, and x7 increased the coefficient of determination to $R^2 = 0.6295$, representing the largest improvement in explanatory power ($\Delta R^2 = 0.292$, $F = 75.945$, $p < 0.001$). This substantial increase can be attributed primarily to the contribution of oil revenues (x7). In the third block (H3), the inclusion of x5, x6, and x10 further increased the coefficient of determination to $R^2 = 0.8154$, yielding a statistically significant improvement ($\Delta R^2 = 0.186$, $F = 48.338$, $p < 0.001$). Since this final model contains the same ten explanatory variables as the full OLS model, its predictive performance is effectively equivalent. However, the methodological value of hierarchical regression lies not in improving prediction accuracy but rather in revealing the cumulative explanatory contribution of each theoretically defined group of variables and assessing their incremental importance in explaining variations in the budget deficit.

3.4 The Machine Learning Models

A. Random Forest: A total of 500 decision trees were constructed using the training dataset. Variable importance was assessed using the %IncMSE (Percentage Increase in Mean Squared Error) measure, and the results are presented in Table (4).

Table (4): Importance of Explanatory Variables Based on the Percentage Increase in Mean Squared Error (%IncMSE) in the Random Forest Model.

Variable	Name	%IncMSE	IncNodePurity	Rank
x7	Oil Revenues	**22.49%**	**4.10e+15**	**1**
x6	Exchange Rate	**12.77%**	**2.00e+15**	**2**
x1	Total Public Debt	**12.35%**	**1.62e+15**	**3**
x2	Actual Expenditures	**11.85%**	**1.43e+15**	**4**
x5	Surplus Reserve	**11.30%**	**1.71e+15**	**5**
x4	Investment	**11.08%**	**1.43e+15**	**6**
x10	Foreign Investments	**10.58%**	**1.47e+15**	**7**
x8	Foreign Reserves	**9.40%**	**1.28e+15**	**8**
x3	Currency Reserves	**7.24%**	**1.10e+15**	**9**
x9	Other Revenues	**5.92%**	**1.41e+15**	**10**

B. XGBoost (eta = 0.05, max_depth = 3, nrounds = 200): Variable importance measured by the Gain metric is presented in Table (5).

Table (5): Variable Importance in the XGBoost Model (Gain).

Variable	Name	Gain	Cover	Frequency	Rank
x7	**Oil Revenues**	**37.71%**	**21.53%**	**14.93%**	**1**
x1	**Total Public Debt**	**11.18%**	**5.09%**	**11.34%**	**2**
x6	**Exchange Rate**	**10.95%**	**11.27%**	**9.26%**	**3**
x2	**Actual Expenditures**	**10.79%**	**15.74%**	**20.93%**	**4**
x9	**Other Revenues**	**7.65%**	**6.76%**	**5.34%**	**5**
x8	**Foreign Reserves**	**7.10%**	**5.44%**	**7.59%**	**6**
x4	**Investment**	**6.86%**	**16.62%**	**12.18%**	**7**
x3	**Currency Reserves**	**5.39%**	**9.21%**	**11.51%**	**8**
x5	**Surplus Reserve**	**2.36%**	**8.34%**	**6.92%**	**9**

The ranking of variable importance is largely consistent across both models, particularly with respect to the top-ranked variable (x7: Oil Revenues) and the other leading variables. This consistency enhances the reliability and robustness of the identified importance rankings. In terms of predictive performance, the Random Forest (RF) model achieved a Test R^2 of 0.8906 and an RMSE of 5655973, while the XGBoost model achieved a Test R^2 of 0.8657 and an RMSE of 5118388. Although XGBoost produced a slightly lower prediction error in terms of RMSE, Random Forest demonstrated superior explanatory and predictive performance as reflected by its higher Test R^2 value.

4. General Comparison of the Models

Table (6): Comparison of the Performance of the Five Models on the Testing Dataset

Model	Adjusted R^2 (In-Sample)	Test R^2 (Out-of-Sample)	RMSE	MAPE%	Rank
Stepwise ★	**0.8042**	**0.9695**	**2,340,941**	**137.19%**	**1**
OLS	**0.8026**	**0.9601**	**2,685,120**	**176.00%**	**2**
Hierarchical	**0.8026**	**0.9601**	**2,685,120**	**176.00%**	**2**
Random Forest	—	**0.8906**	**5,655,973**	**417.92%**	**4**
XGBoost	—	**0.8657**	**5,118,388**	**345.51%**	**5**

The Stepwise Regression model clearly outperformed all other models, achieving the best overall predictive performance. In contrast, the machine learning models exhibited relatively weaker performance. This result can be attributed to the fact that the macroeconomic relationships in Iraq during the study period were characterized by a considerable degree of structural stability and relative linearity, which provided linear regression models with both explanatory and predictive advantages over machine learning algorithms. Moreover, machine learning methods generally require larger sample sizes to fully exploit their ability to capture complex nonlinear patterns. Therefore, their comparatively lower performance in this study should not be interpreted as a weakness of the algorithms themselves, but rather as a consequence of the limited sample size and the nature of the underlying economic relationships, which constrained their ability to generalize effectively. Linear models, on the other hand, benefit from relatively stable macroeconomic structures that can be efficiently captured through linear relationships. It is also worth noting that Adjusted R^2 is not calculated for machine learning models in the traditional statistical sense, and therefore direct comparisons based on this measure are not applicable**.**

5. Consensus Analysis of Variable Importance Across the Models

The use of multiple modeling approaches in this study is not primarily intended to compare performance indicators; rather, it aims to identify the variables that consistently demonstrate importance regardless of the analytical method employed. When a variable is simultaneously highlighted by a classical statistical model such as Stepwise Regression, a structurally grounded Hierarchical Regression model, and a nonlinear Machine Learning model, this convergence provides strong cross-method evidence that is difficult to dismiss. It should also be emphasized that the perspective adopted here is primarily statistical rather than economic. Accordingly, the discussion that follows should be viewed as a statistical interpretation that raises important questions and insights rather than offering definitive economic explanations.

Table (7): Consensus Matrix Across the Four Models

Variabel	Variable Name	Stepwise	Hierarchical	RF (رتبةRF (Rank)	XGBoost (Rank)	Consensus	Category
x7	Oil Revenues	✓ **p<0.001**	**H2-Primary Driver**	**22.49% #1**	**37.71% #1**	**4/4**	**A**
x2	Actual Expenditures	✓ **p<0.001**	**H1** — Core Variable	**11.85% #4**	**10.79% #4**	**4/4**	**A**
x4	Investment	✓ **p<0.001**	**H2** — Significant	**11.08% #6**	**6.86% #7**	**4/4**	**A**
x9	Other Revenues	✓ **p<0.001**	**H1** —Core Variable	**5.92% #10**	**7.65% #5**	**4/4**	**A**
x10	Foreign Investment	✓ **p<0.001**	**H3** — Significant	**10.58% #7**	—	**3/4**	**B**

x6	Exchange Rate	**p=0.136**	**H3 — Significant**	**12.77% #2**	**10.95% #3**	**3/4**	**B**
x8	Foreign Reserves	**p=0.111**	**H1 —Core Variable**	**9.40% #8**	**7.10% #6**	**3/4**	**B**
x1	Public Debt	✗ Excluded	**H1 — Core Variable**	**12.35% #3**	**11.18% #2**	**2/4**	**C**
x5	Surplus Reserve	✗ Excluded	**H3 — Significant**	**11.30% #5**	**2.36% #9**	**2/4**	**C**
x3	Currency Reserves	✗ Excluded	**H2 — Significant**	**7.24% #9**	**5.39% #8**	**1/4**	**D**

5.1 First: Category A – Complete Consensus Across All Four Models (4/4 Models)

Oil revenues (x7) emerged as the most influential variable without dispute, ranking first in both the Random Forest model (22.49%) and the XGBoost model (37.71%), while also being highly significant in the Stepwise Regression model ($p < 0.001$). Moreover, it was the primary factor responsible for the 29.2% increase in the hierarchical model's R^2. The positive coefficient ($\beta > 0$) is particularly intriguing and warrants further investigation by economic specialists to determine whether it reflects specific government spending patterns during periods of oil-revenue booms. Actual expenditures (x2) achieved full consensus across all models and displayed a negative coefficient ($\beta = -0.933$). This raises an important question as to whether the negative relationship is driven by its interaction with x7 or whether the variable captures the degree of budget execution rather than total expenditure, an issue deserving closer examination. Investment (x4) exhibited a negative coefficient ($\beta = -7262$), which is broadly consistent with prior expectations. Of particular interest is the exact equality in absolute magnitude between the coefficients of x4 and x10 (+7262), accompanied by opposite signs. Such a non-random pattern may indicate a structural relationship between domestic and foreign investment that merits further investigation. Finally, other revenues (x9) showed a positive coefficient ($\beta = +1.839$), raising the possibility of reverse causality; that is, these revenues may be collected in response to rising budget deficits rather than acting as their cause. In this regard, a Granger causality test could provide valuable additional evidence.

5.2 Second: Category B – Partial Consensus Across Three of the Four Models (3/4 Models)

Exchange Rate (x6) is the most intriguing variable within this category. Although it is not statistically significant in the linear models ($p = 0.136$), it ranks second in the Random Forest model and third in the XGBoost model. The most plausible statistical explanation for this discrepancy is the presence of a nonlinear effect or an effect that operates only beyond a certain threshold, which can be captured by machine learning algorithms but not adequately represented by conventional linear models. Foreign Reserves (x8) were retained in the Stepwise Regression model according to the AIC criterion despite their relatively weak significance ($p = 0.111$) and appeared in most models with moderate importance. Its close association with oil revenues may partially explain its failure to achieve statistical significance in the linear models. Foreign Investment (x10) was highly significant in the Stepwise Regression model ($p < 0.001$) but did not appear among the top-ranked variables in the XGBoost model. Furthermore, its positive coefficient, combined with the opposite sign observed for x4 (Investment), represents an interesting statistical pattern that warrants further investigation.

5.3 Third: Category C – Methodological Disagreement (2/4 Models)

Public Debt (x1) represents the most notable case of disagreement among the models. It was excluded from the Stepwise Regression model and was not statistically significant in the OLS model ($p = 0.257$), yet it ranked third in the Random Forest model and second in the XGBoost model. This striking discrepancy strongly suggests the presence of a nonlinear relationship. While the linear model searches for a consistent linear effect and fails to detect one, tree-based algorithms are capable of identifying conditional and nonlinear patterns. This raises an important question: does the effect of public debt emerge only after exceeding a certain threshold level? This remains an open issue for further investigation. Surplus Reserve (x5) also exhibits methodological disagreement, even within the machine learning models themselves. It ranked fifth in the Random Forest model (11.30%) but only ninth in the XGBoost model

(2.36%). The substantial difference between these importance measures suggests that the variable's influence is likely driven by its interactions with other explanatory variables rather than by a strong independent effect of its own.

5.4 Fourth: Category D – Marginal Influence (1/4 Consensus)

Currency Reserves (x3) fall into the category of marginal influence. The variable was excluded from the Stepwise Regression model and exhibited one of the weakest levels of statistical significance in the OLS model ($p = 0.702$). The most direct statistical interpretation is that the information contained in this variable is already captured by other, more influential variables included in the model. This does not necessarily imply that currency reserves are economically unimportant; rather, it suggests that they do not provide substantial additional explanatory power once the effects of the remaining variables have been taken into account.

6. Conclusion

The findings of the four models indicated that the oil revenues, actual expenditures, investment and other revenues were found to be important factors in all the modelling methods used in the study. This is consistency and implies that their explanatory power is strong and not tied to a specific modeling approach. This result is logical from an economic point of view because oil production is the main income source for the Iraqi budget, and expenditure and investment by the government are the main instruments of fiscal policy. Meanwhile, non-oil revenues are a complementary source which increases government capacity to deal with fiscal imbalances. The linear estimation yielded a positive coefficient for oil revenues, but it should be treated with caution because it could simply be the result of a time of increased oil revenues coinciding with the increase in public spending. The estimated relationship is therefore more likely to reflect the dominant fiscal conduct during the study period, and is not necessarily a causal effect.

References

1. Adeyinka, O., & Ekum, M. I. (2020). Application of hierarchical polynomial regression models to predict transmission of COVID-19 at global level. International Journal of Clinical Biostatistics and Biometrics, 6(1).
2. Breiman, L. Random forests machine learning 45 (1), 5-32 (2001) 10.1023. A: 1010933404324.
3. Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794).
4. David, B., Ghassan, B. (2015). "Hierarchical Regression for Analysis of Multiple Outcomes", Amercan Journal of Epidemiology VOL. 182, NO. 5.
5. Fox, J. (1991). "Regression Diagnostics: An Introduction". Sag University Paper Series on Quantitative.
6. Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. Annals of statistics, 1189-1232.
7. Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning (2nd ed.). Springer.
8. Henderson, H. V. , Velleman, P. F. ,(1981) "Building Multple Regression Model in Intertvey" Journal of the Biomtrics , 37, 391-411.
9. Huberty, C. J. (1989) "Problems With Stepwise Methods- Better Altenatives. In B. Thompson (ED), Advances in Social Science Methodology" VOL. 1, PP. 43-70 .
10. Kerlinger, F. N. (1986). Foundations of behavioral research (3rd ed.). New York: Holt, Rinehart and Winston.
11. Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). Applied Linear Statistical Models (5th ed.). New York: McGraw-Hill.
12. Lewis,M, (2007) "stepwise versus hierarchical regression :pros and cons "university of texas.
13. Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). Introduction to Linear Regression Analysis (6th ed.). Hoboken, NJ: John Wiley & Sons.
14. Pedhazur, E. J. (1997). Multiple regression in behavioral research (3rd ed.). Orlando, FL: Harcourt Brace. Selvin, H. C., & Stuart, A. (1966). Data-dredging procedures in survey analysis. The American Statistician, 20(3), 20-23.
15. Peter,F (1999) "Hierarchical regression analysis in struchral equation modeling " department of psychology and pedagogics vrije university .
16. Rehman, W. & Fatima, G. (2012) "Consequential Effects of Budget Defection Economic Growth of Pakistan" International Journal of Business and Social Science , VOL 3 , NO 7, April (2012) .
17. Saleh, A. S., & Harvie, C. (2005). The budget deficit and economic performance: A survey. The Singapore Economic Review, 50(02), 211-243.
18. Thompson, P. (1995). " Stepwise Regression and Stepwise Discriminant Analysis Need Not Apply Her : A Guidelines Editorial".
19. Turkson, A. J., & Otchey, J. E. (2015). Hierarchical multiple regression modelling on predictors of behavior and sexual practices at Takoradi Polytechnic, Ghana. Global journal of health science, 7(4), 200.
20. Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. Journal of Applied Econometrics, 16(3), 289–326.