

Determinants of Non-Hydrocarbon Economic Growth in Algeria (1970–2019): Evidence from Shrinkage Estimation Methods.

محددات النمو الاقتصادي خارج قطاع المحروقات في الجزائر (1970–2019):
مقاربة بأساليب تقدير الانكماش.

LAZALI Souhila^{1*}

(Laboratory LAMOPS)

Ecole National Supérieur en Statistique
et d'Economie Appliquée, Algeria.

lazali.souhila@ensea.edu.dz

MEZOUE Fatiha²

(Laboratory LAMOPS)

Ecole National Supérieur en Statistique
et d'Economie Appliquée, Algeria.

mezoued.fatiha@ensea.edu.dz

Received: 16th/04/2026;

Accepted: 26th/05/2026

Abstract:

Algeria's persistent hydrocarbon dependence has structurally constrained non-oil sector development, rendering the identification of non-hydrocarbon growth determinants both analytically critical and policy-relevant. This study applies Ridge, LASSO, and Elastic Net shrinkage estimators to model non-hydrocarbon GDP growth over 1970–2019, addressing severe multicollinearity in macroeconomic time-series data. Structural breaks at 1990 and 2000, identified via the Bai-Perron test, are accommodated through dummy variables. An 80/20 temporal train-test split with 5-fold cross-validation is employed for model selection and predictive evaluation. All shrinkage estimators outperform OLS out-of-sample, with Elastic Net achieving the largest MSE reduction of 49.07%. Labour force participation emerges as the dominant growth determinant, while government expenditure reveals a robust fiscal effect concealed by classical estimation. Domestic savings and inflation are systematically excluded under sparsity constraints. These findings highlight labour force development and productive public spending as primary levers for non-hydrocarbon diversification in Algeria.

Keywords: Economic Growth; Multicollinearity; Shrinkage Estimators; Ridge; LASSO; Elastic Net .

Jel Classification Codes : O40; C52; C51

* Corresponding author: LAZALI Souhila

Abstract in Arabic:

يُشكّل الاعتماد الهيكلي المزمّن على المحروقات في الجزائر عائقاً أمام تطور القطاعات غير النفطية، مما يجعل تحديد محددات النمو خارج المحروقات مسألة ذات أولوية تحليلية وسياساتية بالغة. تستند هذه الدراسة إلى مقدرات الانكماش Ridge و LASSO و Elastic Net لتقدير نموذج الناتج المحلي الإجمالي غير النفطي خلال الفترة 1970-2019، في ظل معالجة دقيقة لإشكالية التعدد الخطي الحاد الملازمة لبيانات السلاسل الزمنية الاقتصادية الكلية. وُظف اختبار Bai-Perron للكشف عن كسرين هيكليين في عامي 1990 و 2000، تمّ استيعابهما من خلال إدراج متغيرات وهمية في جميع المواصفات المقدّرة. اعتمد تقسيم زمني بنسبة 20/80 مقروناً بالتحقق المتقاطع الخماسي لضبط المعاملات وتقييم الأداء التنبؤي. كشفت النتائج عن تفوق جميع مقدرات الانكماش على طريقة المربعات الصغرى العادية خارج العينة، إذ حقّق Elastic Net أعلى انخفاض في متوسط مربع الخطأ بنسبة بلغت 49.07%. يتصدّر مشاركة القوى العاملة قائمة المحددات الأكثر تأثيراً في النمو، في حين يُظهر الإنفاق الحكومي أثراً إيجابياً ومعنوياً ظلّ مُهمّماً في إطار التقدير الكلاسيكي. تُرسّخ هذه النتائج مكانة مقدرات الانكماش بديلاً منهجياً متفوقاً في السياقات الاقتصادية الكلية المعرضة للتعدد الخطي.

Keywords (in Arabic): النمو الاقتصادي; التعدد الخطي; مقدرات الانكماش; Ridge; LASSO; Elastic Net .

Jel Classification Codes : O40; C52; C51

Introduction

Algeria's economic performance has long been dominated by its hydrocarbon sector, rendering overall growth acutely sensitive to fluctuations in international oil prices and leaving non-oil sectors structurally underdeveloped. Despite successive reform and diversification strategies, the country's growth trajectory remains vulnerable to external shocks, underscoring a persistent failure to activate the productive potential of non-hydrocarbon activities. Against this backdrop, the empirical identification of the key macroeconomic determinants of non-hydrocarbon economic growth is not only analytically important but constitutes a prerequisite for the design of effective diversification policies.

Empirical investigations in this context, however, face significant econometric challenges. Macroeconomic time-series data, particularly over long horizons, are frequently characterized by strong multicollinearity among explanatory variables. This phenomenon inflates the variance of OLS estimators, rendering coefficient estimates statistically unstable and potentially misleading, even when the estimator retains its theoretical

property of unbiasedness. The inadequacy of OLS under multicollinearity thus motivates the adoption of alternative estimation strategies capable of delivering more reliable inference and improved predictive performance.

In response to these methodological limitations, the statistical and econometric literature has developed a class of regularization, or shrinkage, estimators. Ridge regression (Hoerl and Kennard, 1970) introduces an L2 penalty that shrinks all coefficients toward zero, trading bias for variance reduction and stabilizing matrix inversion in near-singular designs. LASSO (Tibshirani, 1996) employs an L1 penalty that can shrink some coefficients exactly to zero, simultaneously performing estimation and automatic variable selection. Elastic Net (Zou and Hastie, 2005) combines L1 and L2 penalties to address LASSO's limitations in settings with groups of highly correlated predictors. Despite their growing adoption in finance and macroeconomic forecasting, the application of these shrinkage methods to the analysis of non-hydrocarbon growth determinants in resource-dependent economies remains relatively scarce.

Against this background, this study addresses two interrelated questions: What are the key macroeconomic determinants of non-hydrocarbon economic growth in Algeria? And how effectively do shrinkage estimators, Ridge, LASSO, and Elastic Net enhance estimation accuracy and out-of-sample predictive performance relative to OLS in the presence of multicollinearity?

The study is guided by three empirically testable hypotheses: (i) significant multicollinearity among macroeconomic regressors deteriorates the reliability of OLS estimates; (ii) shrinkage estimators improve coefficient stability and reduce out-of-sample prediction error relative to OLS; and (iii) the Elastic Net estimator, by combining L1 and L2 penalties, delivers superior overall performance through simultaneous variance reduction and variable selection.

2. Theoretical Framework and Literature Review

The empirical growth literature provides a rich theoretical backdrop for the selection of relevant determinants. The neoclassical growth model of Solow (1956) emphasized capital accumulation, labor, and exogenous technological progress as the fundamental engines of long-run growth. Endogenous growth models, notably those of Romer (1986, 1990) and Lucas (1988), subsequently endogenized technological progress by

highlighting the roles of human capital, knowledge spillovers, and increasing returns to scale, thereby expanding the set of policy-relevant growth drivers. Barro (1990) integrated government expenditure into an endogenous growth framework, showing that productive public spending can enhance long-run growth, while Fischer (1993) established macroeconomic stability particularly low and stable inflation as a prerequisite for sustained growth. The roles of foreign direct investment (Borensztein et al., 1998) and trade openness (Sachs and Warner, 1995) have also been extensively examined, with effects found to be conditional on the structural characteristics of the host economy.

Within the Algeria-specific literature, recent empirical studies have employed diverse time-series methodologies to explore growth determinants, often yielding mixed results attributable to the hydrocarbon-dominated economic structure. Toumache and Rouaski (2016), using MICMAC and VCS methods, identified labor force, investment, productivity, and external sector performance as the most critical variables shaping Algeria's future growth prospects, emphasizing that these determinants operate through complex, interconnected mechanisms rather than independently. LeHocine et al. (2025) applied ARDL and ECM frameworks to examine the relationships between GDP per capita growth, FDI, inflation, and unemployment from 2000 to 2023, finding complex short- and long-run dynamics with limited FDI spillovers to non-hydrocarbon sectors. Driss and Ouahrani (2023) similarly found FDI effects on growth to be mediated through investment channels but sensitive to exchange rate volatility, while Madani et al. (2023) documented weak or negative FDI impacts attributable to its concentration in the hydrocarbon sector. Haouas et al. (2021), using a growth accounting framework, identified heavy reliance on factor accumulation rather than total factor productivity as a structural characteristic of Algeria's growth, pointing to the need for efficiency-enhancing reforms in non-hydrocarbon sectors. Hamadouche et al. (2024) further documented bidirectional causality between public expenditure components and growth, while highlighting the simultaneity and endogeneity challenges that complicate fiscal policy inference in hydrocarbon-dependent contexts.

Collectively, the Algeria-specific empirical literature reveals three salient patterns. First, the hydrocarbon sector's dominance systematically distorts standard macroeconomic relationships, rendering growth dynamics in non-oil sectors structurally distinct. Second, the mixed and sometimes

contradictory findings across studies are partly attributable to the econometric challenges inherent in multicollinear macroeconomic time-series data. Third, despite the growing methodological sophistication of the literature, shrinkage estimation methods have not yet been applied to the analysis of non-hydrocarbon growth determinants in Algeria, representing a clear gap that the present study seeks to address.

3. Data, Variables, and Empirical Methodology

3.1 Data and Variable Selection

This study employs annual time-series data covering the period 1970–2019 for Algeria, yielding 50 observations. Data are sourced from two databases: the World Bank's World Development Indicators (WDI) database for the explanatory variables, and the National Office of Statistics (ONS) of Algeria for the dependent variable.

The dependent variable is the Gross Domestic Product excluding the hydrocarbon sector (GDP_NH), used as a direct measure of non-hydrocarbon economic growth. This variable is computed as the difference between total GDP and hydrocarbon GDP, both expressed in constant prices, as reported in the national accounts published by the ONS. The selection of explanatory variables is grounded in the theoretical frameworks reviewed in Section 2, and encompasses six core macroeconomic determinants. Table 1 provides a summary of variable definitions and data sources.

Table 1: Variable Definitions and Units of Measurement

Symbol	Variable	Definition	Unit
GDP_NH	Non-Hydrocarbon GDP	Total GDP minus hydrocarbon sector GDP, measured at constant prices, reflecting the market value of goods and services produced outside the oil and gas sector	Constant DZD (billions)
GEXP	Government Expenditure	General government final consumption expenditure, reflecting public sector spending on goods and services	Constant USD (millions)
FDI	Foreign Direct Investment	Net inflows of foreign direct investment from abroad, capturing cross-border capital flows	Constant USD (millions)
GDS	Gross Domestic Savings	GDP minus total final consumption expenditure, representing the portion of income saved domestically	Constant USD (millions)

Determinants of Non-Hydrocarbon Economic Growth in Algeria (1970–2019): Evidence from Shrinkage Estimation Methods.

/ LAZALI Souhila & MEZOUEF Fatiha

LAB	Labor Force Participation	Total active labor force (ages 15–64) as a share of the working-age population	Number of persons
INF	Inflation Rate	Annual percentage change in the Consumer Price Index (CPI), measuring the rate of price level change	% (annual)
OPEN	Trade Openness	Sum of exports and imports of goods and services expressed as a percentage of GDP	% of GDP

Source: Prepared by the researcher based on ONS national accounts and World Bank WDI.

$$GDP_NH = f(GEXP, FDI, GDS, LAB, INF, OPEN).$$

The general functional form of the growth relationship is expressed as:

A logarithmic transformation is applied to all variables to linearize the growth relationships and allow coefficients to be interpreted as elasticities.

Prior to estimation, the presence and severity of multicollinearity among the explanatory variables is assessed through three complementary diagnostics. The pairwise correlation matrix provides a preliminary indication of linear associations between explanatory variables, though it remains limited to bivariate relationships and fails to capture more complex dependencies involving multiple regressors simultaneously (Kennedy, 2008). A more comprehensive measure is the Condition Number (CN),

computed as : $\kappa = \sqrt{\lambda_{\max} / \lambda_{\min}}$, where $\lambda_{\max}$ and $\lambda_{\min}$ denote the largest and smallest eigenvalues of the standardized $X'X$ matrix, respectively; a value exceeding 30 is indicative of severe multicollinearity (Belsley et al., 1980). Finally, the Variance Inflation Factor (VIF), defined for the j -th regressor

as: $VIF_j = \frac{1}{1 - R_j^2}$ where R_j^2 is the coefficient of determination obtained from regressing X_j on the remaining explanatory variables, provides a variable-specific measure of variance inflation; values exceeding 10 are conventionally regarded as indicative of severe multicollinearity (Kutner et al., 2005), necessitating corrective estimation strategies.

3.2 Shrinkage Estimators

This section presents the theoretical foundations of shrinkage estimators employed in this study.

3.2.1 Ridge Regression

Ridge regression, introduced by Hoerl and Kennard (1970), addresses the problem of multicollinearity by modifying the OLS objective function through the addition of an L_2 norm penalty on the coefficient vector. The

Ridge estimator minimizes:
$$\min_{\beta} \left\{ (y - X\beta)'(y - X\beta) + k \sum_{j=1}^k \beta_j^2 \right\},$$

The resulting estimator is: $\hat{\beta}_{Ridge} = (X'X + kI)^{-1} X'y,$

where ($k \geq 0$) is the shrinkage parameter and I is the identity matrix of order p . By adding a positive constant k to each diagonal element of $X'X$, ridge regression stabilizes matrix inversion even when $X'X$ is near-singular, thereby reducing the variance inflation that characterizes OLS estimates under multicollinearity. The ridge estimator is biased, with bias vector and mean squared error (MSE) given by:

$$\begin{aligned} Bias(\hat{\beta}_{RE}) &= -k(X'X + kI_p)^{-1} \beta, \\ MSE(\hat{\beta}_{RE}) &= \sigma^2 \text{tr} \left[(X'X + kI_p)^{-1} X'X (X'X + kI_p)^{-1} \right] + k^2 \beta' (X'X + kI_p)^{-2} \beta. \end{aligned}$$

This expression exhibits a classic bias–variance trade-off: as k increases, the variance component (first term) decreases while the bias component (second term) increases. The net effect on MSE depends on the magnitude of k relative to the signal-to-noise ratio of the data. Importantly, there exists an optimal value of k for which the MSE of the ridge estimator is strictly lower than that of OLS, even though the former is biased.

3.2.2 LASSO (Least Absolute Shrinkage and Selection Operator).

The Least Absolute Shrinkage and Selection Operator (LASSO), introduced by Tibshirani (1996), extends the ridge regression framework by replacing the L_2 penalty with an L_1 norm penalty on the coefficient vector. The LASSO estimator is obtained by minimizing the following penalized objective function:

$$\min_{\beta} \left\{ (y - X\beta)'(y - X\beta) + \lambda \sum_{j=1}^k |\beta_j| \right\}.$$

where $\lambda \sum_{j=1}^p |\beta_j|$ denotes the L_1 norm of the coefficient vector and $\lambda \geq 0$ is the regularization parameter. Unlike ridge regression, which shrinks all coefficients continuously toward zero without eliminating any, the geometric properties of the L_1 penalty induce exact sparsity in the solution: for sufficiently large values of λ , some coefficients are shrunk precisely to zero, thereby performing simultaneous estimation and automatic variable selection.

Unlike ridge regression, which admits explicit closed-form expressions for both the bias vector and the MSE, the LASSO estimator does not possess a closed-form solution due to the non-differentiability of the L_1 penalty at the origin. Consequently, the bias and MSE of the LASSO estimator are not tractable analytically and must be evaluated empirically.

When two or more regressors are strongly correlated, LASSO tends to arbitrarily select one and discard the others, yielding unstable variable selection and potentially misleading inference (Zou and Hastie, 2005).

3.2.3 Elastic Net

Elastic Net, introduced by Zou and Hastie (2005), was specifically developed to overcome the limitations of both ridge regression and LASSO, particularly in settings characterized by groups of highly correlated predictors. Elastic Net addresses this deficiency by combining both the L_1 and L_2 norm penalties into a unified regularization framework. The Elastic Net estimator is obtained by minimizing the following penalized objective function:

$$\min_{\beta} \left\{ (y - X\beta)'(y - X\beta) + \lambda \left(\alpha \sum_{j=1}^k |\beta_j| + (1 - \alpha) \sum_{j=1}^k \beta_j^2 \right) \right\}.$$

where $\lambda \geq 0$ is the overall regularization parameter controlling the total amount of shrinkage, and $\alpha \in [0, 1]$ is the mixing parameter governing the relative contribution of each penalty. When $\alpha = 1$, Elastic Net reduces exactly to LASSO; when $\alpha = 0$, it reduces to ridge regression. For intermediate values of α , Elastic Net combines the variable selection capability of LASSO with the grouping effect and stability of ridge regression, simultaneously shrinking correlated predictors together rather than arbitrarily discarding some of them (Zou and Hastie, 2005).

As with LASSO, the Elastic Net estimator does not admit a closed-form solution due to the presence of the L_1 penalty component, and its bias and MSE cannot be expressed analytically in closed form (Hastie et al., 2009).

The bias–variance trade-off under Elastic Net is more flexible than that of either ridge regression or LASSO individually. By tuning α , the analyst can continuously interpolate between the two penalty regimes, adapting the estimator to the specific structure of the data (Hastie et al., 2009).

The empirical performance of Elastic Net in macroeconomic forecasting and growth modeling has been documented in several recent studies, including Coulombe et al. (2020), who found that Elastic Net consistently outperforms both ridge and LASSO in terms of out-of-sample predictive accuracy for macroeconomic time-series, and De Mol et al. (2008), who demonstrated the effectiveness of regularization methods in high-dimensional macroeconomic settings.

3.3 Data Splitting and Cross-Validation

To ensure a rigorous and unbiased assessment of out-of-sample predictive performance, the dataset is partitioned into a training set comprising 80% of observations (1970–2009) and a test set comprising the remaining 20% (2010–2019). All models are estimated exclusively on the training set, and predictive accuracy is evaluated on the held-out test set. This temporal split respects the chronological ordering of time-series data and prevents data leakage. Within the training set, 5-fold cross-validation is employed to select the optimal regularization parameters for each shrinkage estimator, balancing model complexity against predictive accuracy.

3.4 Model Evaluation Criteria

Model performance is evaluated using both in-sample and out-of-sample criteria. In-sample fit is assessed through the coefficient of determination R^2 and the Adjusted R^2 , which penalizes for model complexity.

Out-of-sample predictive accuracy is evaluated on the hold-out test set using three complementary metrics. The Mean Squared Error (MSE) measures the average squared deviation between observed and predicted values:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \text{ The Root Mean Squared Error } RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2},$$

$$\text{The Mean Absolute Error (MAE): } MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Shrinkage estimators are preferred over OLS when they achieve lower out-of-sample prediction error and improved coefficient stability.

4 . Results and discussion :

4.1 Descriptive Statistics and Multicollinearity Diagnostics

Before proceeding to estimation, a preliminary examination of the statistical properties of the variables is warranted. Table 2 presents the descriptive statistics of the log-transformed variables used in the analysis.

Table 2: Descriptive Statistics of Log-Transformed Variables

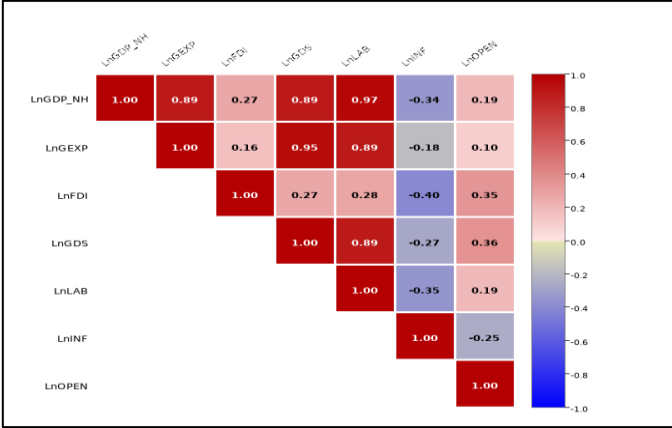
Variable	Mean	Median	SD	Skewness	Kurtosis
LnGDP-NH	13.466	13.877	2.104	−0.240	1.665
LnGEXP	22.958	22.900	1.035	−0.450	2.778
LnFDI	17.260	19.211	5.228	−1.793	5.840
LnGDS	23.734	23.602	1.097	−0.458	2.863
LnLAB	16.537	16.589	0.431	−0.240	1.618
LnINF	1.834	1.782	0.867	−0.475	4.124
LnOPEN	4.046	4.066	0.194	−0.684	3.165

Source: Prepared by the researcher using R (version 4.4.1)

Table 2 presents the descriptive statistics of all logarithmic variables over the study period 1970–2019. The results show that LnFDI exhibits the highest variability (SD = 5.228) and negative skewness (−1.793), reflecting the volatility of foreign direct investment flows into Algeria. All other variables display moderate skewness, suggesting approximate normality in their distributions .

Figure 1 presents the Pearson correlation matrix of log-transformed variables reveals several noteworthy relationships.

Figure 1: Pearson Correlation Matrix of Log-Transformed Variables



Source: Prepared by the researcher using R (version 4.4.1)

Figure 1 reveals that LnGDP_NH is strongly correlated with LnLAB ($r = 0.97$), LnGDS ($r = 0.89$), and LnGEXP ($r = 0.89$), while a high inter-predictor correlation between LnGEXP and LnGDS ($r = 0.95$) signals potential multicollinearity. LnINF exhibits negative associations across most variables, notably with LnFDI ($r = -0.40$), and LnOPEN shows consistently weak correlations ($r \leq 0.36$), suggesting a limited role in non-hydrocarbon growth.

The eigenvalues of the standardized correlation matrix, in descending order, are $\lambda_1 = 4.0656$, $\lambda_2 = 1.3929$, $\lambda_3 = 0.7853$, $\lambda_4 = 0.5774$, $\lambda_5 = 0.1645$, $\lambda_6 = 0.0127$, and $\lambda_7 = 0.0017$, yielding a condition number of $\kappa = \sqrt{2459.473} = 49.59$, substantially exceeding the critical threshold of 30 established by Belsley et al. (1980).

Table 3 presents the Variance Inflation Factors (VIF) for all model regressors.

Table 3: Variance Inflation Factors (VIF) for Model Variables

	ln(GEXP)	ln(GDS)	ln(LAB)	ln(OPEN)	ln(INF)	ln(FDI)
VIF	38.2037	41.0215	6.4682	3.6743	1.4458	1.3464

Source: Prepared by the researcher using R (version 4.4.1)

Table 3 confirms severe multicollinearity among the explanatory variables, led by ln(GDS) and ln(GEXP), whose tolerance levels fall far below acceptable limits. These diagnostic results confirm the presence of severe multicollinearity in the model, providing strong justification for the

adoption of shrinkage-based estimation methods, namely Ridge, LASSO, and Elastic Net regression.

4.2 Stationarity Tests

Prior to estimation, the stationarity properties of the series were assessed using the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests. As reported in Appendix A, most log-transformed variables are non-stationary in levels and integrated of order one, $I(1)$, which is consistent with the typical behavior of macroeconomic time series.

While conventional OLS estimation of non-stationary series risks yielding spurious results, first-differencing—the standard remedy—eliminates valuable long-run information. Given that the primary objective of this study is predictive accuracy rather than structural inference, the strict stationarity requirement was not imposed (Seghiri, 2020).

4.3 Structural Break

To account for potential structural instability and parameter shifts within the time-series continuum, the Bai-Perron (1998) algorithm for multiple unknown structural breaks was implemented.

Table 4 : Bai-Perron Structural Break Test: Model Selection Criteria

m (Breakpoints)	RSS	BIC	ΔBIC	Optimal
m = 0	0.7289	-38.222	—	
m = 1	0.2200	-66.815	-28.593	
m = 2	0.0627	-98.257	-31.442	✓
m = 3	0.0354	-95.539	+2.718	

Source: Prepared by the researcher using R (version 4.4.1)

Table 5 : Identified Structural Break Dates and Sub-periods (Algeria, 1970–2019)

Period	Years	Observations	Economic Context
Period 1	1970 – 1990	21	Pre-structural adjustment era
Period 2	1991 – 2000	10	IMF structural adjustment & reforms
Period 3	2001 – 2019	19	Post-reform & oil boom era

Source: Prepared by the researcher using R (version 4.4.1)

As detailed in Table 4, the model selection metrics conclusively favor a two-break configuration ($m=2$). This specification minimizes the Bayesian Information Criterion (BIC) at a global minimum of -98.257 , exhibiting a substantial improvement over the single-break model ($\Delta BIC=-31.442$), whereas expanding the framework to a three-break alternative ($m=3$) yields a deteriorated BIC of -95.539 .

The identified optimal breakpoints segment the 1970–2019 temporal horizon into three distinct macro-economic regimes (Table 5). To safeguard against parameter bias arising from these structural regime shifts, two temporal dummy variables are integrated into the baseline framework to absorb the level shifts. Consequently, the baseline econometric model is specified as follows:

$$\ln(\text{GDP_NH}_t) = \beta_0 + \sum_{i=1}^k \beta_i X_{i,t} + \delta_1 D_{1990,t} + \delta_2 D_{2000,t} + \varepsilon_t$$

where β_0 denotes the intercept; β_i represents the elasticity coefficients associated with the vector of macroeconomic predictors $(X_{\{i,t\}})$; and δ_1 and δ_2 capture the level shifts stemming from the structural breaks identified in 1990 and 2000, respectively. The temporal dummies are defined conditionally such that for $D_{1990,t}=1$, and $D_{2000,t}=1$ (0 otherwise). ε_t is the error term assumed to be independently and identically distributed as $\varepsilon_t \sim N(0, \sigma^2)$, and t indexes the annual time period from 1970 to 2019.

4.4 Estimation Results: OLS and Shrinkage Estimators

Table 6 reports the coefficient estimates, shrinkage parameters, goodness-of-fit measures, and predictive accuracy metrics for each model.

Table 6 : Comparison of OLS, Ridge, LASSO, and Elastic Net Regression.

Variable	OLS	Ridge	LASSO	Elastic Net
<i>Coefficient Estimates</i>				
Intercept	-70.8661*** (0.000)	-28.3584*** (0.000)	-68.2315*** (0.000)	-58.8498*** (0.000)
ln(GEXP)	0.1046 (0.467)	0.4097*** (0.000)	0.0375. (0.054)	0.1717*** (0.000)
ln(FDI)	-0.0036 (0.378)	0.0010 (0.876)	-0.0033 (0.226)	-0.0003 (0.384)
ln(GDS)	-0.0807 (0.461)	0.2295*** (0.000)	0.0000 (Excl.)	0.0000 (Excl.)
ln(LAB)	4.9826*** (0.000)	1.6887*** (0.000)	4.8437*** (0.000)	4.1055*** (0.000)
ln(INF)	-0.0150 (0.671)	-0.0767. (0.096)	0.0000 (Excl.)	-0.0069 (0.240)
ln(OPEN)	0.3965* (0.046)	-0.4222*** (0.000)	0.1990 (0.104)	0.0706 (0.234)

Determinants of Non-Hydrocarbon Economic Growth in Algeria (1970–2019): Evidence from Shrinkage Estimation Methods.

/ LAZALI Souhila & MEZOUEF Fatiha

D1 (1990)	0.0509 (0.742)	1.1768*** (0.000)	0.0782 (0.328)	0.3585*** (0.000)
D2 (2000)	-0.2416* (0.030)	0.4216*** (0.000)	-0.1459* (0.034)	0.0000 (Excl.)
Shrinkage Parameters				
Lambda (λ)	—	0.2079	0.0026	0.0130
λ_{1se}	—	0.4377	0.0730	0.0632
L1 Ratio (α)	—	0.00	1.00	0.50
Active Vars. (p)	8	8	6	6
Goodness-of-Fit				
R ²	0.9971	0.9887	0.9970	0.9964
Adjusted R ²	0.9965	0.9865	0.9965	0.9959
F-statistic	1745.47*** (0.000)	511.84*** (0.000)	2753.39*** (0.000)	2698.71*** (0.000)
Predictive Accuracy				
MSE (Train)	0.0140	0.0560	0.0204	0.0143
RMSE (Train)	0.1184	0.2365	0.1429	0.1194
CV-MSE (5-Fold)	—	0.0686	0.0191	0.0237
MSE (Test)	0.0161	0.0143	0.0093	0.0082
RMSE (Test)	0.1268	0.1194	0.0962	0.0906
MAE (Test)	0.1151	0.0834	0.0807	0.0748
Information Criteria				
AIC	-56.39	12.26	-60.58	-51.37
BIC	37.27	29.47	-47.20	-37.99

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. P-values are reported in parentheses.

Source: Prepared by the researcher using R (version 4.4.1)

4.5 Discussion of Individual Coefficient Estimates

The empirical results across OLS, Ridge, LASSO, and Elastic Net models reveal both strong consistency in core structural relationships and meaningful differences arising from regularization. Overall model fit is exceptionally high in all specifications ($R^2 \approx 0.988$ – 0.997), indicating strong explanatory power; however, predictive performance and coefficient stability provide clearer distinctions between approaches than in-sample fit alone.

A key robust finding is the strong positive and statistically significant effect of labor ($\ln(\text{LAB})$) across all models. The magnitude remains large in OLS

(4.98), Ridge (1.69), LASSO (4.84), and Elastic Net (4.11), confirming labor as the dominant driver of the dependent variable. This stability across shrinkage methods suggests that labor is not sensitive to multicollinearity or overfitting concerns and likely represents a structurally fundamental factor in the underlying economic relationship.

Government expenditure ($\ln(\text{GEXP})$) exhibits model-dependent behavior. While insignificant in OLS, it becomes highly significant and positive in Ridge and Elastic Net, and marginally significant in LASSO. This shift indicates that multicollinearity or omitted variable noise in the classical specification may have suppressed its effect. Regularization appears to stabilize its contribution, suggesting that fiscal activity has a more consistent positive role than OLS alone implies.

Foreign direct investment ($\ln(\text{FDI})$) remains statistically insignificant across all models, with coefficients near zero. This consistency suggests that FDI does not exert a direct measurable impact in this specification, or its effect may be indirect, mediated through other macroeconomic channels not captured here.

Gross domestic savings ($\ln(\text{GDS})$) is excluded by LASSO and Elastic Net, while showing no significance in OLS and a positive effect in Ridge. Its exclusion under sparsity constraints implies that its explanatory contribution is weak once penalization is applied, reinforcing the conclusion that savings is not a key determinant in the presence of other macroeconomic controls.

Inflation ($\ln(\text{INF})$) is consistently insignificant or weakly negative, and is also excluded in LASSO. This indicates a lack of stable relationship in the current model structure, suggesting either nonlinear effects or context-specific transmission mechanisms not captured in a linear log-log framework.

Trade openness ($\ln(\text{OPEN})$) shows mixed results: positive and weakly significant in OLS, negative and strongly significant in Ridge, and insignificant in LASSO and Elastic Net. This sign instability signals sensitivity to model specification and suggests that the effect of openness is not robust; it may depend on interaction effects or structural breaks not explicitly modeled.

Dummy variables capture structural shifts differently across estimators. D1 (1990) becomes strongly significant and positive under Ridge and Elastic Net, suggesting that regularization reveals structural change effects that are not captured in OLS. D2 (2000) is negative and significant in OLS and LASSO but positive in Ridge and insignificant in Elastic Net, indicating instability in interpreting this structural period across estimation methods.

From a methodological perspective, shrinkage parameters highlight model complexity control. Ridge ($\alpha = 0$) retains all variables with coefficient shrinkage, while LASSO ($\alpha = 1$) enforces sparsity, removing $\ln(\text{GDS})$ and $\ln(\text{INF})$. Elastic Net ($\alpha = 0.5$) balances both approaches, yielding a parsimonious model with strong predictive performance. Notably, Elastic Net achieves the lowest test RMSE (0.0906) and MAE (0.0748), outperforming all alternatives in out-of-sample accuracy despite slightly lower R^2 .

Information criteria further support penalized models, with LASSO achieving the lowest AIC (-60.58) and competitive BIC (-47.20), indicating superior trade-off between fit and complexity. Although OLS exhibits the highest R^2 , it is less competitive in predictive evaluation, highlighting overfitting risk in unregularized estimation.

In summary, the results suggest a clear hierarchy: labor is the most robust determinant; government expenditure gains credibility under regularization; while FDI, savings, and inflation show limited explanatory power. Elastic Net emerges as the most balanced estimator, combining strong predictive accuracy with model parsimony and improved stability.

4.6 Out-of-Sample Predictive Performance

Table 7 presents the out-of-sample predictive accuracy of the four estimated models on the held-out test sample (20% of observations). All three regularised estimators outperform OLS across every metric, confirming the overfitting limitation of unpenalised classical regression in the presence of severe multicollinearity.

Table 7 : Out-of-Sample Predictive Accuracy Comparison — Test Sample (20%)

Metric	OLS	Ridge	LASSO	Elastic Net
<i>MSE (Test)</i>	0.0161	0.0143	0.0093	0.0082
<i>RMSE (Test)</i>	0.1268	0.1194	0.0962	0.0906
<i>MAE (Test)</i>	0.1151	0.0834	0.0807	0.0748
<i>MAPE % (Test)</i>	0.9087	0.6116	0.5800	0.5302
Reduction vs. OLS	—	-11.18%	-42.24%	-49.07%

Note. Reduction vs. OLS = (Model MSE – OLS MSE) / OLS MSE × 100. Bold = best metric .

Source: Prepared by the researcher using R (version 4.4.1)

Elastic Net achieves the greatest predictive gains, reducing MSE by 49.07% relative to OLS (test MSE = 0.0082, RMSE = 0.0906, MAE = 0.0748, MAPE = 0.5302%), followed by LASSO with a 42.24% MSE reduction (test MSE = 0.0093, RMSE = 0.0962), and Ridge with an 11.18% improvement (test MSE = 0.0143). These results unambiguously establish Elastic Net as the superior estimation strategy, balancing variable selection, coefficient shrinkage, and out-of-sample generalisation capacity for non-hydrocarbon GDP growth modelling in Algeria.

Conclusion

This study employed Ridge, LASSO, and Elastic Net shrinkage estimators to identify the macroeconomic determinants of non-hydrocarbon economic growth in Algeria over the period 1970–2019, addressing the severe multicollinearity confirmed by a condition number of 49.59 and VIF values exceeding 41.02 for certain regressors. The inclusion of structural break dummy variables for 1990 and 2000, identified through the Bai-Perron test, further enhanced the empirical robustness of the estimated models.

The empirical findings consistently identify labour force participation as the dominant determinant of non-hydrocarbon growth across all specifications, while government expenditure emerges as a robust and positive contributor under Ridge and Elastic Net, revealing a significant fiscal growth effect that classical OLS estimation fails to capture. Domestic savings and inflation are systematically excluded under sparsity-inducing constraints, suggesting limited direct explanatory power within this framework. Among the competing estimators, Elastic Net delivers superior out-of-sample predictive performance, reducing test-set MSE by 49.07% relative to OLS, thereby validating the adoption of shrinkage methods in macroeconomic modelling contexts characterised by high predictor correlation.

From a policy standpoint, the prominence of labour force participation as the principal growth driver underscores the critical importance of human capital development, vocational training, and labour market reforms as foundational pillars of Algeria's non-hydrocarbon diversification strategy. The consistent positive effect of government expenditure further highlights the role of productive public spending in stimulating non-oil sector activity, provided that fiscal resources are channelled toward growth-enhancing investment rather than consumption subsidies.

Several limitations of the present study merit acknowledgment. The sample of 50 annual observations, while adequate for classical econometric analysis, represents a relatively small dataset for machine learning methods and may constrain the stability of cross-validation procedures. Furthermore, potential endogeneity between non-hydrocarbon GDP and key regressors, notably government expenditure and labour force participation, is not addressed through instrumental variable methods, which may introduce bias in the estimated elasticities. The omission of institutional quality indicators and human capital measures, attributable to data availability constraints, may additionally result in omitted variable bias.

These limitations delineate several productive avenues for future research. Extending the present framework to a panel of resource-dependent economies would enhance the external validity of the findings and permit the examination of cross-country heterogeneity in the labour-growth nexus. Incorporating time-varying parameter models or Markov-switching specifications could more precisely capture the structural shifts induced by hydrocarbon price cycles. Finally, integrating shrinkage estimation within a cointegration framework, such as ARDL or VECM, would provide deeper insight into the long-run properties of the relationships modelled and strengthen the econometric foundations of the analysis.

References:

- [1]. Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55–67.
- [2]. Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B*, 58(1), 267–288.
- [3]. Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B*, 67(2), 301–320.
- [4]. Solow, R. M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70(1), 65–94.
- [5]. Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, 94(5), 1002–1037.
- [6]. Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, 98(5), S71–S102.
- [7]. Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22(1), 3–42.

- [8]. Barro, R. J. (1990). Government spending in a simple model of endogenous growth. *Journal of Political Economy*, 98(5), S103–S125.
- [9]. Fischer, S. (1993). The role of macroeconomic factors in growth. *Journal of Monetary Economics*, 32(3), 485–512.
- [10]. Borensztein, E., De Gregorio, J., & Lee, J.-W. (1998). How does foreign direct investment affect economic growth? *Journal of International Economics*, 45(1), 115–135.
- [11]. Sachs, J. D., & Warner, A. M. (1995). Natural resource abundance and economic growth. NBER Working Paper No. 5398.
- [12]. Toumache, R., & Rouaski, K. (2016). Prospective analysis of Algerian economic growth by 2025: A structural analysis. *Journal of Applied Business Research*, 32(3), 791–804.
- [13]. LeHocine, A., Benbouziane, M., & Ali, A. (2025). Exploring the drivers of economic growth in Algeria: An ARDL and ECM analysis of unemployment, FDI, and inflation. *South Florida Journal of Development*, 6(2).
- [14]. Driss, A., & Ouahrani, A. (2023). The impact of foreign direct investment on economic growth: Empirical evidence from Algeria (1990–2021). *Dafatir Iqtisadiya Journal*.
- [15]. Madani, H., Medjenah, F., & Yessaad, A. (2023). Assessment of the impact of foreign capital flows on economic growth in Algeria. *Journal of Innovations and Sustainability*, 7(3), Article 8.
- [16]. Haouas, A., Ochi, A., & Labidi, M. (2021). Sources of Algeria's economic growth, 1979–2019: Augmented growth accounting framework and growth regression method. *Regional Science Policy & Practice*, 13(6), 1860–1878.
- [17]. Hamadouche, F., Hamadouche, B., Ghandour, M., Hamadouche, A., & Boutrig, B. (2024). Granger test and ARDL model for socio-economic challenges: Public expenditure components, economic growth, and inflation. *SocioEconomic Challenges*, 8(4), 70–87.
- [18]. Kennedy, P. (2008). *A Guide to Econometrics* (6th ed.). Blackwell Publishing.
- [19]. Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. Wiley.
- [20]. Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied Linear Statistical Models* (5th ed.). McGraw-Hill.
- [21]. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning* (2nd ed.). Springer.

Determinants of Non-Hydrocarbon Economic Growth in Algeria (1970–2019): Evidence from Shrinkage Estimation Methods.

/ LAZALI Souhila & MEZOUE Fatiha

- [22]. Hastie, T., Tibshirani, R., & Wainwright, M. (2015). Statistical Learning with Sparsity: The LASSO and Generalizations. CRC Press.
- [23]. Coulombe, P. G., Leroux, M., Stevanovic, D., & Surprenant, S. (2020). How is machine learning useful for macroeconomic forecasting? *Journal of Applied Econometrics*, 37(5), 920–964.
- [24]. De Mol, C., Giannone, D., & Reichlin, L. (2008). Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components? *Journal of Econometrics*, 146(2), 318–328.
- [25]. Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 66(1), 47–78
- [26]. Seghiri, S. (2020). Determinants of private investment in Algeria for the period (2002-2017) *Economic Sciences, Management And Commercial Science Review*, 12(3), 209-222.

Appendices

Appendix A: Unit Root Test Results for Log-Transformed Variables

Variable	ADF Test		PP Test		Conclusion
	Statistic	p-value	Statistic	p-value	
LnGDP-NH	-0.6324	0.9702	-0.1956	0.9900	I(1)
LnGEXP	-2.8656	0.2271	-6.2808	0.7422	I(1)
LnFDI	-2.7438	0.2759	-35.1562	0.0100	I(0)/I(1)
LnGDS	-2.4519	0.3930	-7.5298	0.6656	I(1)
LnLAB	-3.0325	0.1601	3.9132	0.9900	I(1)
LnINF	-2.8838	0.2198	-18.3109	0.0738	I(0)/I(1)
LnOPEN	-1.9306	0.6021	-8.8224	0.5863	I(1)

Source: Prepared by the researcher using R (version 4.4.1)