Comparing Mathematical Models for Forecasting the Youth Unemployment Rate in Jordan

Fawaz Sharaf

Received May 30, 2024
Accepted August 14, 2024
Electronic access September 15, 2024

Background: Unemployment is one of the numerous socioeconomic challenges that exist in all countries around the world. Governments and various stakeholders rely on accurate forecasting of unemployment rates to make informed decisions relating to investments and fiscal and monetary policies. Method: The main objective of this study is to compare and examine the predictive accuracy of multiple linear regression (MLR), time series autoregressive distributed lag (ARDL) and vector autoregression (VAR) models for predicting the youth unemployment rate of Jordan. These models are chosen because they are established econometric models, each grounded in specific theoretical principles and practical applications. Using time series annual data from 1991 to 2022, the three models will be fitted to model the youth unemployment rate of Jordan based on (1) population growth rate (annual %), (2) current gross domestic product (GDP) in US$, (3) foreign direct investment net (BoP current US$), (4) external debt stocks as a percentage of gross national income (% of GNI), (5) foreign direct investment net inflows (% of GDP), (6) foreign direct investment net outflows (% of GDP), (7) literacy rate adult total (% of people ages 15 and above), (8) school enrollment tertiary (gross), gender parity index (GPI) and (9) inflation consumer prices (annual %). The predictive accuracy of the models will be assessed using mean absolute error, root mean square error, and mean absolute percentage error. Results: The findings of the study will show that the ARDL model has the least mean absolute error, root mean square error and mean absolute percentage error, hence, having the highest predictive accuracy. Conclusion: Due to limitations associated with sample size, time series period and selected lag values, the researcher notes the current findings while suggesting further refinement in the methodology for future studies.

Introduction

Background and Context

Unemployment is a critical socioeconomic issue that affects every nation globally. It significantly impacts the living standards and socioeconomic status of individuals. Consequently, the unemployment rate is a critical metric for governments and policymakers as it forms the basis for developing strategies to address unemployment[1]. Accurate forecasting of unemployment rates is crucial for effective economic planning and policy-making. This study focuses on Jordan, where youth unemployment is a significant challenge, impacting economic stability and growth. Various models, including MLR, ARDL and VAR, have been developed to forecast unemployment rates, each with distinct advantages and limitations. The primary cause of rising unemployment rates is often inadequate demand within the economy to support sufficient employment opportunities[2]. During periods of reduced demand, businesses require less labor, leading to reduced work hours or layoffs. Unemployment is driven by fundamental economic changes and is influenced by frictional, structural and cyclical factors[3]. Frictional unemployment occurs due to the time delays inherent in finding new employment in a free market. Structural unemployment arises from factors such as skill mismatches between workers and job requirements or geographical barriers preventing relocation to areas with available jobs[4]. Additionally, inadequate wage rates can lead to a lack of motivation to seek employment, contributing to a skills gap despite available opportunities. Technological advancements, increased competition and government policies also contribute to structural unemployment. Cyclical unemployment results from the cyclical nature of economic indicators within a business cycle[2]. It decreases as the economy expands and optimizes output. Conversely, a decline in gross domestic product (GDP) leads to an economic downturn and increased cyclical unemployment. Researchers worldwide continually develop and review models to enhance the accuracy of unemployment forecasts[5]. This study aims to compare the predictive accuracy of three models for forecasting youth unemployment in Jordan: MLR, ARDL and VAR.

Problem Statement and Rationale

The study aims to identify the most accurate model for forecasting the youth unemployment rate in Jordan. Previous studies have shown varying performance of different models, necessitating a comprehensive comparative analysis to determine the best approach for this context.

Significance and Purpose

This research provides a comparative analysis of three forecasting models (MLR, ARDL and VAR) to determine which model offers the highest predictive accuracy for Jordan's youth unemployment rate. The findings will aid policymakers and stakeholders in making informed decisions based on reliable economic forecasts.

Objectives

- To compare the predictive accuracies of MLR, ARDL and VAR models for the youth unemployment rate in Jordan.
- To evaluate the models using statistical parameters such as MAE, RMSE and MAPE.
- To provide recommendations for future research and model improvements.

Scope and Limitations

The study focuses on time series data from 1991 to 2022, covering various economic indicators relevant to unemployment forecasting. Limitations include potential biases in historical data, the chosen time period, and the specific lag values used in the models.

Theoretical Framework

This study leverages econometric modeling techniques to analyze and forecast unemployment rates, providing a framework for comparing the performance of different models. MLR, ARDL and VAR were chosen because they are established econometric models, each grounded in specific theoretical principles and practical applications.

Literature Review

Youth Unemployment

Youth unemployment is a significant challenge that affects both developed and developing countries. The challenge is especially acute in developing countries due to the prevalence of extreme poverty, which often compels all household members to work to ensure survival[6]. The issue of youth unemployment poses a significant threat to society across various dimensions, including social, economic and political spheres, particularly in many developing nations. From a social perspective, youth unemployment concerns not only the unemployed individuals but also society at large. Many young individuals expect to secure employment upon completing their education. However, unemployment can lead to demoralization, a decrease in their skills and knowledge, and a decline in their chances of finding a job, ultimately leading to social exclusion[7]. Economically, youth unemployment has negative consequences, including labor market instability, higher living costs, reduced tax revenue, and wasted investments in education and training. Politically, youth unemployment is often seen as a catalyst for both political stability and conflicts within nations. Selim et al.[8] argue that youth unemployment is a primary factor in political instability and conflicts. Researchers have attributed popular uprisings in various Arab countries, including Jordan, to the significant issue of youth unemployment[9]. In Middle Eastern countries, increasing unemployment is primarily attributed to the unbalanced relationship between economic growth and rapid population growth, resulting in a persistent increase in the unemployment rate, as exemplified by Jordan[10]. The unemployment situation in Jordan is particularly alarming, given the large proportion of youth in the total population. Nearly two-thirds of Jordan's population is under the age of 30, and annually, 32% of youth between the ages of 15 and 30 are unable to find employment. This equates to over 100,000 young individuals in Jordan entering the job market each year. The situation has worsened in recent years due to the country's implementation of a complete lockdown in response to the COVID-19 pandemic since March 2020, leading to a significant increase in the unemployment rate[8]. High levels of unemployment in Jordan indicate the economy's failure to effectively utilize its labor resources. Various factors contribute to unemployment, including low levels of general economic activity, recession, inflation, population growth, rapid technological changes, disability, willingness to work, and discrimination, among others.

Modeling and Forecasting Macroeconomic Variables

Modelling and forecasting of macroeconomic variables are essential for analyzing various aspects of a country's economic condition. Numerous studies have employed diverse time series models to predict macroeconomic variables. The unemployment rate is a significant macroeconomic variable, and its modeling and forecasting are crucial for making various economic decisions[11]. Various models have been employed for modeling and forecasting, as well as for comparing the forecasting performance of models concerning the unemployment rate in multiple countries. Unemployment rates are typically studied using econometric models, which involve analyzing stationary time series, seasonality, trend analysis and exponential smoothing. These models often incorporate the VAR, Autoregressive Integrated Moving Average (ARIMA) technique, ARDL and Ordinary Least Squares (OLS) method, among others. The VAR model is an important time series model that effectively captures the dynamic patterns of macroeconomic variables and facilitates accurate forecasting. According to Uylangco & Li[12], the accuracy of forecasts made with VAR models can be examined by measuring the trace of the mean-squared forecast error matrix or the generalized forecast error second moment.[13] demonstrated higher precision in forecasting certain macroeconomic variables using VAR models compared to alternative models such as transfer functions. Therefore, it can be argued that utilizing combined forecasts derived from VAR models is an effective approach for enhancing the accuracy of predictions.

The ARDL model is an econometric model used to analyze the relationships between various time series variables over short and long periods. In the ARDL model, the autoregressive component represents the past values of the dependent variable. Jmaii[14] found that ARDL models could accurately model economic time series data in examining the influence of the health crisis on the volatility of commodity price returns. Similarly, Shrestha & Bhatta[15] noted that since ARDL models account for the impact of both lagged values of dependent and independent variables, they are significant in modeling economic time series as economic factors are widely dependent on their past and present occurrences. This explains why ARDL models are widely considered superior to least squares estimations in econometrics[11].

Methodology

This section presents the methodology employed in conducting a comparative analysis of models used to forecast the unemployment rate. Specifically, it discusses the data used, economic variables included in the models, models fitted, and model evaluation metrics.

Study Design

The study employed both descriptive and correlational research designs. Descriptive research design seeks to explain the current status of an identified variable, while correlational research seeks to determine the extent of the relationship between two or more variables. The descriptive research design was used in the description and comparison of individual models for forecasting the unemployment rate in Jordan, while correlational research design was used to examine the relationship between the youth unemployment rate in Jordan and indicators such as population growth rate and GDP.

Variables

To determine the best model to forecast the Jordanian unemployment rate, annual unemployment data covering the period from 1991 to 2022 was used. The data was downloaded from the World Bank Open Data website[16]. The study consisted of ten variables: one dependent variable, the youth unemployment rate, and nine independent variables. The choice of independent variables was determined by findings from existing literature on the determinants of unemployment[17],[2],[18],[19],[10],[6].

(i) Dependent Variable

Youth Unemployment Rate: The percentage of the labor force aged between 15 and 24 that is without work but available for and seeking employment[6]. This metric is crucial for understanding the employment challenges faced by the youth in Jordan.

(ii) Independent Variables

Population growth rate (annual %): The annual percentage increase in the population. A higher population growth rate can lead to an increased supply of labor, potentially affecting the unemployment rate if job creation does not keep pace.

Current GDP (US$): The total monetary value of all goods and services produced within a country. It serves as an indicator of economic health; higher GDP generally indicates better economic conditions, which can reduce unemployment.

Foreign direct investment (net, BoP, current US$): The net inflows of investment intended to acquire a lasting management interest (10 percent or more of voting stock) in a business operating in an economy other than that of the investor. It includes equity capital, reinvestment of earnings and other forms of long-term and short-term capital. This investment can influence economic growth and job creation, thereby potentially reducing unemployment.

External debt stocks (% of GNI): Includes debt owed to non-residents repayable in currency, goods, or services. Total external debt comprises public, publicly guaranteed, and private nonguaranteed long-term debt, use of International Monetary Fund (IMF) credit, and short-term debt, which includes all debt with an original maturity of one year or less and interest in arrears on long-term debt. GNI refers to the sum of value added by all resident producers, plus any product taxes (less subsidies) not included in the valuation of output, and net receipts of primary income, compensation of employees, and property income from abroad. High levels of external debt can burden an economy, potentially leading to reduced public and private investment in job creation, thereby impacting economic stability and unemployment. Data for this metric were sourced from the World Bank[16].

Foreign direct investment, net inflows (% of GDP): The amount of investment that flows into an economy to acquire a significant ownership stake (10 percent or more of voting stock) in a company operating in that economy. This term refers to the aggregate of equity capital, reinvestment of earnings, other long-term capital, and short-term capital, as indicated in the balance of payments. Net inflows are calculated as the difference between new investment inflows and disinvestment from foreign investors. These net inflows are then divided by the GDP to express the total value of foreign investments coming into the reporting economy as a percentage of GDP. Increased FDI inflows typically indicate economic growth and job creation, which can reduce unemployment rates. Data for this metric were obtained from the IMF's Balance of Payments database[20], with additional information from the United Nations Conference on Trade and Development United Nations Conference on Trade and Development (UNCTAD)[21].

Foreign direct investment, net outflows (% of GDP): This metric measures the total value of investments made by residents of the reporting economy into businesses abroad, expressed as a percentage of GDP. Direct investment involves acquiring a significant ownership stake (10 percent or more of voting stock) in a business located in another country, indicating a lasting management interest. Net outflows include equity capital, reinvestment of earnings, other long-term capital, and short-term capital. High outflows may suggest strong economic strength, indicating that the domestic economy is robust enough to invest abroad. However, it could also imply less domestic investment, which might negatively impact local job creation depending on the economic context. Data for this metric were sourced from the[20], with additional information from the UNCTAD[21].

Literacy rate, adult Total (% of people ages 15 and above): The percentage of people aged 15 and above who can read and write with understanding. Higher literacy rates generally enhance employability and can reduce unemployment rates by creating a more skilled labor force.

School enrollment, tertiary (gross), gender parity index (GPI): The gender parity index (GPI) for gross enrollment ratio in tertiary education refers to the ratio of women to men enrolled at the tertiary level in public and private schools. The calculation of this indicator involves dividing the gross enrollment ratio of females in tertiary education by the gross enrollment ratio of males in tertiary education.

Inflation, consumer prices (annual %): The annual percentage change in the cost to the average consumer of acquiring a basket of goods and services. High inflation can reduce purchasing power and economic stability, potentially leading to higher unemployment rates.

Data

(a) Data Collection and Transformation

Data were collected from reliable sources including the World Bank[16] and Jordan's Department of Statistics[22]. The data were cleaned and preprocessed to ensure accuracy and consistency. Missing data were handled using appropriate interpolation techniques. To prepare the data for analysis, a log transformation was applied to stabilize variance and normalize the distribution. Negative values were addressed by adding a constant value to the dataset, based on the smallest negative value, as recommended by West[23]. This adjustment ensured all values were positive, allowing for a valid log transformation. Because coefficients on the natural logarithm (LN) scale are directly interpretable as approximate proportional differences, the LN of all variables were taken before applying models. Logging converts exponential growth patterns into linear growth patterns while converting multiplicative seasonal patterns into additive seasonal patterns[24]. Before applying log transformation, the negative values were handled by adding a constant value to the time series data as recommended by West[23]. This is necessary as the logarithm of a negative number is undefined. The constant value added was chosen based on the smallest negative value in the dataset to ensure all transformed values remain positive.

First, we identified the smallest (most negative) value in the dataset. To make all values positive, we added a constant to each value in the dataset. This ensures that the smallest value in the dataset becomes slightly great-er than zero, which is necessary for log transformation. Once all values were adjusted to be positive, we applied the LN transformation to stabilize the variance and normalize the distribution. This process converts exponen-tial growth patterns into linear growth patterns and multiplicative seasonal patterns into additive seasonal pat-terns, making the data more suitable for linear modeling techniques. The adjustment for negative values is a common practice in time series analysis to ensure that all data points can be log-transformed without encoun-tering mathematical issues.

Model Application and Evaluation
The processed data were analyzed using MLR, ARDL and VAR models. Parameters were estimated using R and Stata. Model predictive accuracy was assessed using MAE, RMSE and MAPE. Additionally, confidence intervals and significance levels were calculated to ensure the robustness of the results.

Ethical Considerations
The study adhered to ethical guidelines for data use, ensuring confidentiality and proper attribution of sources. All data were anonymized to protect the privacy of individuals.

Comparison of Models

(a) MLR

The MLR model is chosen for its capability to include multiple independent variables, which allows it to capture the diverse influences on youth unemployment. This model quantifies the relationship between each variable and the youth unemployment rate, providing policymakers with insights into the impact of different factors. The high R-squared value of the MLR model indicates strong predictive power, making it a valuable tool for forecasting and policy-making.

A multiple linear regression is a statistical model that uses more than one independent variable to predict the outcome of a dependent variable. This model facilitates the determination of how much each independent variable contributes to the variation in the dependent variable[25]. In the existing literature, multiple linear regression has been extensively used to predict various economic variables. For instance, Syla[25] used a multiple linear regression to predict employment rates while Monari et al.[26] fitted a multiple linear regression to model economic determinants of youth unemployment and predict future youth unemployment levels in Kenya. Ashfahani et al.[27] highlighted that linear regression is one of the most common models used in economics and econometrics to study the relationship between one or more variables. A multiple linear regression is generally denoted by

$$y = \beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n + \varepsilon$$

where:

- y represents the dependent variable to be predicted,
- X_i represents the independent variables,
- n represents the number of independent variables in the model,
- β_i represents the coefficients of the independent variables,
- β_0 is a constant, and
- ε refers to the error term.

Accordingly, in the context of this study, the fitted regression model was presented as:

$$\begin{aligned}\ln(\text{UN}) = \beta_0 + \beta_1 \ln(\text{POP}) + \beta_2 \ln(\text{GDP}) + \beta_3 \ln(\text{FDI}) \\ + \beta_4 \ln(\text{EXD}) + \ln(\text{Inflow\_FDI}) + \ln(\text{Outflow\_FDI}) \\ + \ln(\text{Lit}) + \ln(\text{GPI}) + \ln(\text{CPI}) + \varepsilon\end{aligned}$$

(b) VAR

VAR is a forecasting method that can be used when two or more time series interact. The selection of the model was informed by economic theories and previous studies. For instance, Aqil et al.[3] used the VAR model to estimate the determinants of unemployment in Pakistan while Alrabba[6] also fitted a VAR model to assess the determinants of the youth unemployment rate in Jordan. Generally, VAR is denoted by

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1n} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{n1} & \varphi_{n2} & \cdots & \varphi_{nn} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \vdots \\ y_{n,t-1} \end{bmatrix} + \cdots +$$

$$\begin{bmatrix} \varphi_{1,t-p} & \varphi_{12,t-p} & \cdots & \varphi_{1n,t-p} \\ \varphi_{2,t-p} & \varphi_{22,t-p} & \cdots & \varphi_{2n,t-p} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{n,t-p} & \varphi_{n2,t-p} & \cdots & \varphi_{nn,t-p} \end{bmatrix} \begin{bmatrix} y_{1,t-p} \\ y_{2,t-p} \\ \vdots \\ y_{n,t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Where, $y_{i,t}$ is the dependent variable
c is a constant (intercept),
i,j is the coefficient of lags of y till order p while
i is the error term
p is the number lages included in the model

VAR models are commonly employed in the modelling of multivariate time series. The structure of the model is such that every variable is a linear function of past lags of itself and past lags of other variables[28]. In the context of this study, this means that Jordan's youth unemployment rate is a linear function of its past lags, and past lags of population growth rate (annual %), current GDP in US$, foreign direct investment, net (BoP, current US$), external debt stocks (% of GNI), foreign direct investment, net inflows (% of GDP), foreign direct investment, net outflows (% of GDP), literacy rate, adult total (% of people ages 15 and above), school enrollment, tertiary (gross), gender parity index (GPI), and inflation, consumer prices (annual %). In econometrics and statistics, lagging for time series data implies that the regression equation used to predict current values of dependent variables is dependent on both current values of the explanatory variables and dependent variable, and past period (lagged) values of the dependent and explanatory variables. For example, suppose that a time series denoted by $x_(t,1)$ is being measured, a VAR model of order 1, denoted by VAR(1) is defined as:

$$x_{t,1} = \alpha_1 + \varphi_{11} x_{t-1,1} + \varphi_{12} x_{t-1,2} + \varphi_{13} x_{t-1,3} + w_{t,1}$$

That is, each variable is a linear function of the lag 1 values for all variables in the set. In the present study, VAR(2) was fitted. The choice of lag was determined by the assessment of the unemployment time series. As shown in Figure 1, a sharp change in the unemployment rate in Jordan seem to be observed every 2 years.

(c) ADRL

Assuming that both dependent and independent variables are lagged up to p and q respectively, ARDL is generally represented by

$$y_t = \delta + \theta_t y_{t-1} + \cdots + \theta_t y_{t-p} + \delta_0 x_t + \delta_1 x_{t-1} + \cdots + \delta_q x_{t-q} + \varepsilon_t$$

where:

- y_t is the dependent variable,
- x_t is the independent variable,
- δ is the impact multiplier,
- θ_t is the distributed lag weight of x_t,

- ε_t is the error term,
- p is the lag length of y_t, and
- q is the lag length of x_t.

The choice of p and q was facilitated by the Akaike information criterion (AIC) and Schwartz information criterion (SIC)[29]. That is, choosing the value of p and q that minimizes the Sum of Squared Errors (SSE) subject to an increase in the number of parameters. Increasing lag lengths increases the number of parameters but reduces SSE. In this study, the choice of lag was limited by the time sample size (number of years) since the regressors cannot be larger than the sample size when fitting the ARDL model. $p = q = 1$, the functional form of the model is denoted by:

Unemployment Rate (UN) = f(GDP, Population Growth Rate (POP), FDI, External Debt Stock (EXD)) + ⋯

In the Distributed Lag (DL) form, the model is represented by:

$$\ln(\mathrm{UN})_t = \beta_0 + \beta_1 \Delta \ln(\mathrm{POP})_t + \beta_2 \Delta \ln(\mathrm{GDP})_t + \beta_3 \Delta \ln(\mathrm{FDI})_t + \beta_4 \Delta \ln(\mathrm{ED}_t) + \cdots + \varepsilon_t$$

In the ARDL form, which includes previous values as part of the independent variables, the model is represented by:

$$\ln(\mathrm{UN}_t) = \beta_0 + \beta_1 \Delta \ln(\mathrm{UN}_{t-1}) + \beta_2 \Delta \ln(\mathrm{POP})_t + \beta_3 \Delta \ln(\mathrm{GDP}_t) + \beta_4 \Delta \ln(\mathrm{FDI})_t + \beta_5 \Delta \ln(\mathrm{ED}_t) + \cdots + \varepsilon_t$$

Lagging every variable up to $t - j$, the model is represented by:

$$\ln(\mathrm{UN}_t) = \beta_0 + \beta_1 \Delta \ln(\mathrm{UN}_{t-j}) + \beta_2 \Delta \ln(\mathrm{POP})_{t-j} + \beta_3 \Delta \ln(\mathrm{GDP}_{t-j}) + \beta_4 \Delta \ln(\mathrm{FDI})_{t-j} + \beta_5 \Delta \ln(\mathrm{ED}_{t-j}) + \cdots + \varepsilon_{t-j}$$

As noted by Handoyo et al.[30], although VAR and ARDL models are similar, their significant difference emanates from the fact that VAR models are used for examining the dynamic interrelationship among multiple variables overtime while ARDL models are employed in the assessment of long-run and short-run dynamics between multiple variables; especially when dealing with non-stationary data. Accordingly, to account for the possibility of the two scenarios in the time series, VAR and ARDL models were evaluated.

Evaluation Metrics

Besides testing the statistical significance of the models at a 5% level of significance and examining the coefficient of determination, the forecasting accuracy of the models was assessed and compared using mean absolute error, root means square error and mean absolute percentage error.

(a) MAE

MAE is a measure of errors between paired observations that express the same phenomenon. MAE can be defined as the average variance between the significant values in the dataset and the projected values in the same dataset[31]. Accordingly, the lower the MAE, the higher the accuracy of the forecasting model.

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{n}$$

(b) RMSE

RMSE is the square root of Mean Square Error (MSE). MSE is the sum of the squared difference between fitted values and observed values divided by the number of historical points less the number of parameters in the model.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

The lower the value of the Root Mean Squared Error, the better the model is.

(c) MAPE

The MAPE measures the average magnitude of error generated by a model[32]. This indicates the average deviation of predictions from the actual values. For example, a MAPE value of 25% indicates that the average absolute percentage deviation between the predictions and the actual values is 25%.

$$\mathrm{MAPE} = \frac{100}{n}\sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

Similar to RMSE and MAE, lower values of MAPE indicate higher accuracy, while higher values indicate lower accuracy.

Analysis, Findings and Discussion

The present study sought to compare and determine the most accurate model for forecasting the Jordanian unemployment rate. The three models evaluated were multiple linear regression, VAR (2) and ARDL (1). In this chapter, the findings of the study are presented and discussed against the findings of the existing literature.

Summary Statistics

As shown in Table 1 below, between 1991 and 2022, the mean youth unemployment rate of Jordan was 32.79% (SD =4.15%) with the highest reported unemployment rate being 42.35% while the lowest unemployment rate over the period being 27.47İThe higher unemployment rate was reported in 2020.

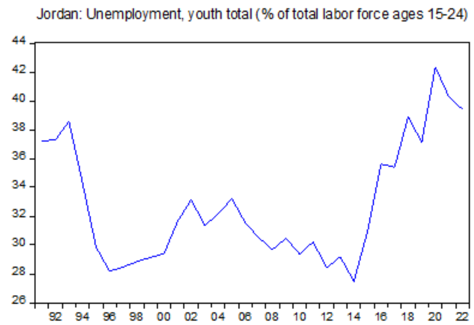

Fig. 1 Unemployment Rate

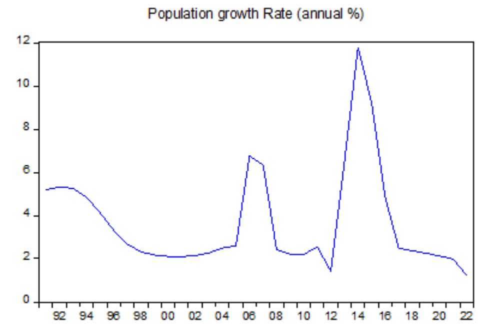

Fig. 2 Population Growth Rate

The findings of the study seem to support the results reported by Antipova[7]. While investigating the impact of COVID-19 on employment and unemployment across multi-dimensional socially disadvantaged areas, Antipova[7] found that the majority of the country reported the highest unemployment rate at the peak of COVID-19. This can be explained by the fact that during COVID-19 lockdowns, the number of non-commuters increased severalfold in 2020 compared to previous years. The increase resulted in significantly reduced activity in several sectors of the economy such as transport hence rendering some people without jobs, Besides, temporary leaves with/without pay, and working from home due to the structural nature of lockdowns rendered some people unemployed[33]. Key statistics for other variables are summarized in Table 1 below.

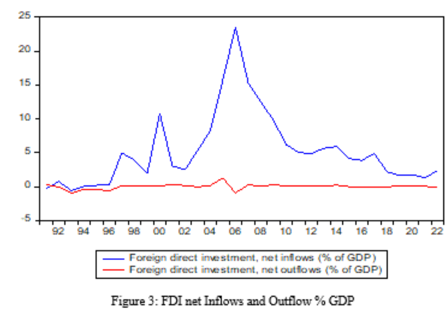

Fig. 3 FDI net Inflows and Outflow % GDP

Indicators

In consistency with the statistics in Table 1, the unemployment time series graph in Figure 1 shows that the lowest youth unemployment rate was reported in 2014 while the highest youth unemployment rate was reported in 2020. Throughout the period, the unemployment rate significantly fluctuated. The most significant drop was reported between 1994 and 1996 while the most significant increase in unemployment was reported between 2014 and 2020.

World Bank. (n.d.). Unemployment, total (% of total labor force)[34]

Similar to the unemployment rate, as illustrated in Figure 2, the population growth rate of Jordan fluctuated throughout the period. The lowest population growth rate was reported in 2012 while the highest rate was reported in 2014 (11.79

World Bank. (n.d.). Population growth (annual %).[34]

Another key observation in the time series is the significant difference between FDI net inflows and FDI net outflows as a percentage of GDP as demonstrated in Figure 3. This is consistent with the findings of Bekhet & Al-Smadi[35] that although FDI inflow performance in Jordan is below other levant regions countries, it is significantly higher than its FDI outflow performance.

World Bank. (n.d.). Foreign direct investment, net inflows (% of GDP). .[34]

As indicated in Figure 4, the total external debt stocks to gross national ratio of Jordan assumed a downtrend between 1991 and 2022. However, between 2012 and 2022, the value rose from 58.56% in 2012 to 94% in 2022.

World Bank. (n.d.). Foreign direct investment, net inflows (% of GDP).[34]

Jordan's inflation rate generally assumed a downward trend throughout the period. However, there were significant

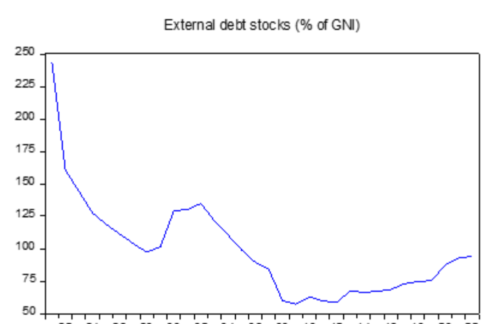

Fig. 4 External debt stocks (% of GNI)

Table 1: Summary Statistics

	CPI	EXD	FDI	GDP__CURRENT_US$_	GPI	INFLOW_FDI	LIT	OUTFLOW_FDI	POP	UN
Mean	17.30027	99.08060	2.62E+09	2.20E+10	1.089902	6.247806	92.94515	2.016259	3.676095	32.79994
Median	17.32028	93.45553	2.86E+09	1.61E+10	1.079215	5.032673	91.72574	2.050427	2.489817	31.44000
Maximum	27.97123	243.6161	3.71E+09	4.75E+10	1.253420	24.53729	98.42000	3.296271	11.79402	42.34900
Minimum	13.12315	57.27757	0.891400	4.34E+09	0.959160	0.401603	85.89232	1.054035	1.226637	27.46800
Std. Dev.	2.897765	38.53452	9.82E+08	1.53E+10	0.079026	5.419445	3.341158	0.388072	2.391648	4.153998
Skewness	1.477211	1.718667	-0.763418	0.372105	0.440044	1.615255	-0.17066	0.066714	1.714528	0.701056
Kurtosis	7.000810	7.200323	2.790598	1.507285	2.239458	5.561381	2.205432	6.692804	5.746027	2.264158
Observations	32	32	32	32	32	32	32	32	32	32

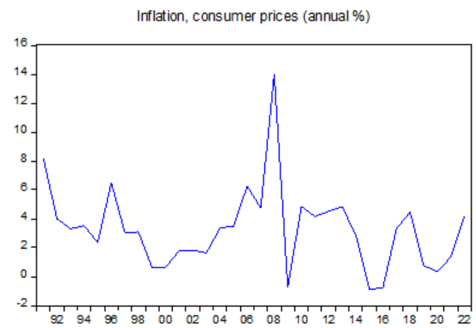

Fig. 5 Inflation, consumer prices (annual %)

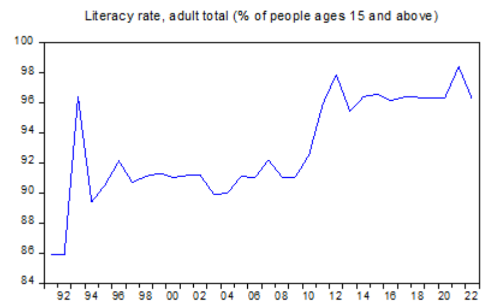

Fig. 7 School enrollment, tertiary (gross), gender parity index (GPI)

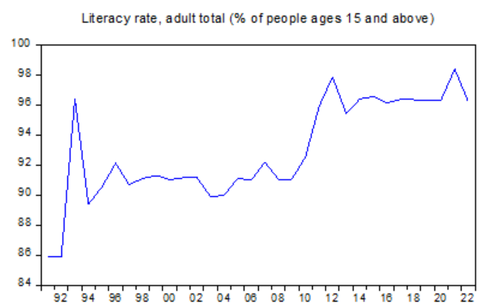

Fig. 6 Literacy rate, adult total (% of people ages 15 and above)

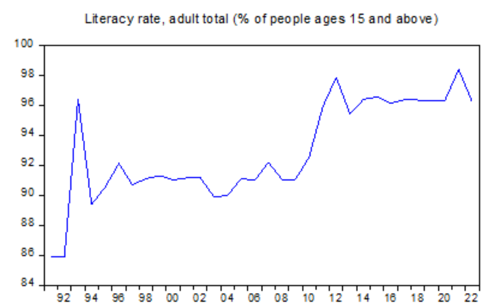

Fig. 8 Current GDP in US$ and Foreign Direct Investment

fluctuations (Figure 5) throughout the period. The highest inflation rate was reported in 2008 (27.97%). The high inflation rate coincided with the global economic crisis of 2008[36].

World Bank. (n.d.). Inflation, consumer prices (annual %).[34]

The Jordan literacy rate was on the rise for the period 1991-2022 (Figure 6). Weldali[37] explains that the rising Jordan literacy rate can be attributed to Jordan's focused efforts on the importance of education awareness and the creation of more literate societies.

World Bank. (n.d.) Literacy rate, adult total (% of people ages 15 and above).[16]

This is further reflected in the increase in school enrollment, tertiary (gross), and gender parity index (GPI) as shown in Figure 7.

World Bank. (n.d.). School enrollment, tertiary (% female). World Bank.[34]

As shown in Figure 8, although the current GDP of Jordan assumed a consistent increase from 1991 to 2022, the net FDI of the country barely changed.

World Bank. (n.d.). Foreign direct investment, net inflows (% of GDP).[34]

Comparing Models

Model 1: MLR

As illustrated by the negative coefficients of several key variables, a more detailed examination reveals how each factor: population growth rate (LN-POP), foreign direct investment net outflows (LN-OUTFLOW-FDI), literacy rate (LN-LIT), foreign direct investment net inflows (LN-INFLOW-FDI), foreign direct investment (LN-FDI) and inflation rate (LN-CPI), negatively influences the unemployment rate in Jordan:

Population growth rate (LN-POP): A negative coefficient indicates that higher population growth rates are associated with lower youth unemployment rates. Although this might seem counterintuitive it could be explained by factors such as increased economic activity and job creation in response to a growing population.

Foreign direct investment net outflows (% of GDP) (LN-OUTFLOW-FDI): The negative coefficient suggests that higher net outflows of FDI are associated with lower youth unemployment rates. This could be due to the benefits of global economic integration and the potential for domestic firms to gain competitive advantages through international operations.

Literacy rate (LN-LIT): A negative coefficient for literacy rate implies that higher literacy rates are associated with lower youth unemployment rates. This is consistent with the expectation that a more literate population is better equipped for employment.

Foreign direct investment net inflows (% of GDP) (LN-INFLOW-FDI): The negative coefficient indicates that higher FDI inflows are linked to reduced youth unemployment. FDI can stimulate economic growth and job creation by bringing in capital, technology, and expertise.

Foreign direct investment (LN-FDI): Similarly, a negative coefficient for overall FDI suggests that increased foreign investment is associated with lower unemployment rates, reflecting the positive impact of foreign capital on job creation.

Inflation rate (LN-CPI): A negative coefficient for the inflation rate implies that higher inflation is associated with lower youth unemployment rates. This relationship can be explained by the short-term trade-off between inflation and unemployment, known as the Phillips curve, where moderate inflation may coincide with economic growth and job creation.

As demonstrated by the p-value ($p < 0.000003$) being less than the 5% level of significance, a conclusion is made that there is significant statistical evidence to indicate that POP, EXT, GDP, FDI, inflow FDI, outflow FDI, Lit, GPI, and CPI can be used to explain the variability in youth unemployment at the 5% level of significance. The R-squared of the model was found to be 0.816509, thus implying that POP, EXT, GDP, FDI, inflow FDI, outflow FDI, Lit, GPI, and CPI explain 81.65% of the variability in youth unemployment in Jordan.

The findings of the model contradict the arguments of Azolibe et al.[17] and Riaz & Zafar[4] that an increase in population, inflation rate and foreign direct investment net outflow results in an increase in the unemployment rate. However, in regards to the population growth rate, Riaz & Zafar[4] argue that the correlation relationship between population growth rate and unemployment depends on the overall economic performance. When economic activity is, a higher population results in higher products and

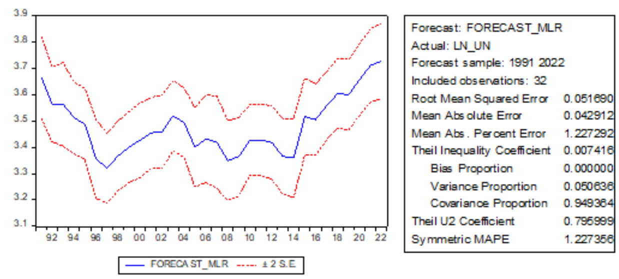

Fig. 9 MLR Model Evaluation

the working population is absorbed hence negating the possible increase in unemployment. Similarly, Agarwal[38] intimate that the relationship between FDI and unemployment is one-sided. For example, in a case where an investing firm produces one product, FDI will result in export substitution thus leading to the home country's unemployment rate increasing. On the contrary, if investing firms are producing more than a single product, FDI to produce one product may result in exports of other products due to the export promotional impact of the foreign affiliate[38]. Similarly, in regards to GPI, EXD and GDP, the findings of the present study contradict the arguments presented in the existing literature that school enrollment, tertiary (gross), gender parity index (GPI), external debt stocks (% of GNI) and current GDP is negatively correlated with the unemployment rate. For example, Alawad et al.[10] posit that a rapid increase in GDP results in a reduction in unemployment. In regards to school enrollment, tertiary (gross), and gender parity index (GPI), Barbu et al.[39] explain the findings of the present study. Barbu et al.[39] assert that if the rate of economic activity is low, an increase in the school enrollment in tertiary education increases the total labour without increasing the opportunity for the people to be absorbed in the job market hence resulting in a higher unemployment rate.

Figure 9: shows that the fitted MLR has an RMSE, MAE and MAPE of 0.05169, 0.0429 and 1.2273,

Model 2: VAR

Similar to the MLR model, it was found that the VAR model is statistically significant in predicting youth unemployment in Jordan. The VAR model reported a higher coefficient of determination (R-squared) of 0.9065. This indicates that the VAR model, with a lag of 2, explains 90.65% of the variability in youth unemployment in Jordan using the following variables: population growth rate (annual %), current GDP in US$, foreign direct investment, net (BoP, current US$), external debt stocks (% of GNI), foreign direct investment, net inflows (% of GDP), foreign direct investment, net outflows (% of GDP), literacy rate, adult total (% of people ages 15 and above), school enrollment, tertiary (gross), gender parity index (GPI), and

Table 2: MLR Output

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LN_POP	-0.003594	0.025506	-0.140891	0.8892
LN_OUTFLOW_FDI	-0.004127	0.077602	-0.053181	0.9581
LN_LIT	-0.588679	0.707579	-0.831963	0.4144
LN_INFLOW_FDI	-0.03205	0.018559	-1.726894	0.0982
LN_GPI	1.2789	0.409303	3.124583	0.0049
LN_GDP	0.068346	0.061473	1.1118	0.2782
LN_FDI	-0.000215	0.003391	-0.063392	0.95
LN_EXD	0.282986	0.074273	3.810094	0.001
LN_CPI	-0.023609	0.078998	-0.298862	0.7679
C	3.27748	2.914544	1.124526	0.2729
R-squared	0.816509	Mean dependent var		3.482989
Adjusted R-squared	0.741445	S.D. dependent var		0.1226
S.E. of regression	0.06234	Akaike info criterion (AIC)		-2.462123
Sum squared resid	0.085498	Schwarz info. criterion (SIC)		-2.00408
Log-likelihood	49.39396	Hannan-Quinn criter.		-2.310295
F-statistic	10.87745	Durbin-Watson stat		1.840901
Prob(F-statistic)	0.000003			

inflation, consumer prices (annual %).

Why Use the VAR Model Multivariate Analysis: The VAR model captures the dynamic relationship between multiple time series variables, treating all variables as endogenous and allowing for the examination of their mutual influences over time.

Lagged Effects: Including lagged values of the variables helps capture temporal dependencies and delayed effects in the data, which is particularly useful in economic forecasting.

Model Fit and Predictive Power: The high R-squared value of 0.9065 indicates that the VAR model fits the data very well and has strong predictive power, making it a valuable tool for forecasting youth unemployment in Jordan with high accuracy.

Understanding Dynamic Interrelationships: By considering the interdependencies among the variables, the VAR model provides a comprehensive understanding of how different economic indicators interact and collectively influence youth unemployment. This holistic approach is crucial for policymakers to design effective interventions.

Breakdown of the VAR Model Components

Population growth rate (annual %):
A high population growth rate can increase the labor supply, which might influence unemployment rates depending on the economic context and capacity to absorb the workforce.

Current GDP in US$:
GDP is a key indicator of economic performance. Variations in GDP reflect the overall health of the economy and its ability to create jobs.

Foreign direct investment (FDI), net (BoP, current US$):
Net FDI measures the total inflow and outflow of foreign investment, which can stimulate economic growth and job creation, thereby impacting unemployment rates.

External debt stocks (% of GNI):
External debt influences government spending and investment, affecting economic growth and employment opportunities.

FDI net inflows and outflows (% of GDP):
These variables measure the direct impact of foreign investment on the economy, which has significant implications for job creation and unemployment.

Literacy rate, adult Total (% of people ages 15 and above):
Literacy rates reflect the education level of the labor force. Higher literacy rates generally enhance employability and can reduce unemployment. School Enrollment, tertiary (gross), gender parity index (GPI): These indicators measure access to higher education and gender equality in education, both of which influence labor market outcomes.

Inflation, consumer prices (annual %):
Inflation affects purchasing power and economic stability. Moderate inflation can coincide with economic growth and job creation, while high inflation can have the opposite effect.

Interpretation of Results

The VAR model, with its high coefficient of determination (0.9065), demonstrates a strong fit to the data, indicating that 90.65% of the variability in youth unemployment can be explained by the selected variables. The dynamic interrelationship among the variables, captured through the lag structure, provides deeper insights into how economic factors influence youth unemployment over time.

Similar to the MLR model, the findings for the VAR model provide valuable information for policymakers. Understanding the lagged effects and interdependencies among the variables can help in designing targeted interventions to reduce youth unemployment in Jordan.

The detailed statistics in Table 3 indicate the overall performance of the VAR model. The high R-squared and significant F-statistic demonstrate that the model is well-suited for explaining the variability in the data. The AIC and SIC values are used for model selection, with lower values indicating a better model fit. The standard error of the equation (S.E. equation) and sum of squared residuals (Sum sq. resids) provide additional measures of the model's accuracy and precision. The model's mean and standard deviation of the dependent variable (Mean dependent and S.D. dependent) are also provided, offering a benchmark for assessing the variability explained by the model. These statistics collectively validate the effectiveness of the VAR model in forecasting youth unemployment in Jordan.

Table 3 VAR Output

R-squared	0.906508
Adj. R-squared	0.698748
Sum sq. resids	0.039934
S.E. equation	0.066611
F-statistic	4.363251
Log-likelihood	56.75790
Akaike AIC	-2.38386
Schwarz SIC	-1.403022
Mean dependent	3.473997
S.D. dependent	0.121362

The generalized fitted VAR equation is:

$$
\begin{aligned}
\text{LN\_UN} = {} & C(1,1) \times \text{LN\_UN(-1)} + C(1,2) \times \text{LN\_UN(-2)} \\
& + C(1,3) \times \text{LN\_POP(-1)} + C(1,4) \times \text{LN\_POP(-2)} \\
& + C(1,5) \times \text{LN\_OUTFLOW\_FDI(-1)} + C(1,6) \times \\
& \text{LN\_OUTFLOW\_FDI(-2)} + C(1,7) \times \text{LN\_LIT(-1)} \\
& + C(1,8) \times \text{LN\_LIT(-2)} + C(1,9) \times \\
& \text{LN\_INFLOW\_FDI(-1)} + C(1,10) \times \\
& \text{LN\_INFLOW\_FDI(-2)} + C(1,11) \\
& \times \text{LN\_GPI(-1)} + C(1,12) \times \text{LN\_GPI(-2)} \\
& + C(1,13) \times \text{LN\_GDP(-1)} + C(1,14) \times \text{LN\_GDP(-2)} \\
& + C(1,15) \times \text{LN\_FDI(-1)} + C(1,16) \times \text{LN\_FDI(-2)} \\
& + C(1,17) \times \text{LN\_EXD(-1)} + C(1,18) \times \text{LN\_EXD(-2)} \\
& + C(1,19) \times \text{LN\_CPI(-1)} + C(1,20) \times \text{LN\_CPI(-2)} \\
& + C(1,21)
\end{aligned}
$$

The estimated VAR equation for youth unemployment in Jordan is:

$$
\begin{aligned}
\text{LN\_UN} = {} & 0.288020556136 \times \text{LN\_UN(-1)} + 0.0348569720807 \times \\
& \text{LN\_UN(-2)} - 0.0574228631537 \times \text{LN\_POP(-1)} + \\
& 0.103112686268 \times \text{LN\_POP(-2)} + 0.01247040398 \\
& \times \text{LN\_OUTFLOW\_FDI(-1)} + 0.0601344910716 \\
& \times \text{LN\_OUTFLOW\_FDI(-2)} - 1.00321910146 \\
& \times \text{LN\_LIT(-1)} - 0.37218148949 \times \text{LN\_LIT(-2)} \\
& - 0.00378053604724 \times \text{LN\_INFLOW\_FDI(-1)} \\
& + 0.00913073640205 \times \text{LN\_INFLOW\_FDI(-2)} \\
& + 0.311088507073 \times \text{LN\_GPI(-1)} - 0.173677273772 \\
& \times \text{LN\_GPI(-2)} - 0.19853359077 \times \text{LN\_GDP(-1)} \\
& + 0.417200777068 \times \text{LN\_GDP(-2)} + 0.00363216704877 \\
& \times \text{LN\_FDI(-1)} + 0.00122271289627 \times \text{LN\_FDI(-2)} \\
& + 0.400494271834 \times \text{LN\_EXD(-1)} + 0.0129800936954 \\
& \times \text{LN\_EXD(-2)} + 0.0883550125946 \times \text{LN\_CPI(-1)} \\
& + 0.109471386397 \times \text{LN\_CPI(-2)} + 0.82104065222
\end{aligned}
$$

The evaluation metrics of the model are shown in Table 4.

The findings of the study show that the VAR model has higher RMSE, MAE, and MAPE values compared to the MLR model. Accordingly, the prediction accuracy of the VAR model is lower than that of the MLR model, as a lower error indicates higher prediction accuracy. However, as indicated by the coefficient of determination, the VAR model explains a higher proportion of variability (90.65%) in the unemployment rate compared to the MLR model (81.65%).

Additionally, Hamilton[40] explains that a higher R-squared value does not necessarily indicate a better fit for the model. A

Table 4: VAR Evaluation Table

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
LN_CPI	32	0.147769	0.113042	3.996624	0.026110
LN_EXD	32	0.186889	0.153458	3.435268	0.020743
LN_FDI	32	3.824964	1.558534	7.675643	0.090562
LN_GDP	32	0.119802	0.102104	0.430880	0.002536
LN_GPI	32	0.026975	0.019478	79.70094	0.121591
LN_INFLOW_FDI	32	0.469442	0.375384	26.77209	0.137083
LN_LIT	32	0.013165	0.010139	0.223380	0.001451
LN_OUTFLOW_FDI	32	0.156643	0.078202	12.44105	0.111459
LN_POP	32	0.535797	0.395136	37.30500	0.230923
LN_UN	**32**	**0.079113**	**0.057881**	**1.653532**	**0.011356**

model with a low R-squared value can still be a good model, while a model with a high R-squared value might be a poor fit for the data. This is because R-squared is a biased estimate of the population R-squared[40].

Model 3: ARDL

The ARDL model was also found to be statistically significant (p-value = 0.000011). The model had the second-highest R-squared value of 0.870869, indicating that it explains 87.07% of the variability in youth unemployment in Jordan. This suggests that the ARDL model, with variables such as population growth rate (annual %), current GDP in US$, foreign direct investment (net, BoP, current US$), external debt stocks (% of GNI), foreign direct investment (net inflows and outflows, % of GDP), literacy rate (adult total, % of people ages 15 and above), school enrollment (tertiary, gross), gender parity index (GPI), and inflation (consumer prices, annual %), provides a robust explanation for the youth unemployment rate in Jordan.

Why Use the ARDL Model

Flexibility in Lag Selection: The ARDL model allows for different lag lengths for each variable, making it more flexible in capturing the dynamics of the data. This flexibility is particularly useful in time series data where the influence of past values can vary across variables.

Handling Non-Stationary Data:

The ARDL model can be applied to data regardless of whether the underlying variables are I(0) or I(1), which means it can handle both stationary and non-stationary time series without the need for differencing. This makes it a versatile tool in econometric modeling.

Short-run and Long-run Relationships:

ARDL models are capable of distinguishing between short-run and long-run relationships among variables. This dual focus provides a comprehensive understanding of how changes in one variable impact others over different time horizons.

Interpretation of Results

Lagged dependent variable (LN-UN(-1)):

The coefficient of the lagged dependent variable (0.368389) indicates the persistence of unemployment over time. Although not statistically significant (p=0.1341), it suggests that past unemployment rates have a moderate impact on current unemployment rates.

Population growth rate (LN-POP):

The positive coefficient (0.021292) suggests a direct but insignificant impact of population growth on youth unemployment, indicating that higher population growth rates could lead to higher youth unemployment.

Foreign direct investment, net inflows and outflows (LN-INFLOW-FDI, LN-OUTFLOW-FDI):

The coefficients for FDI inflows and outflows are mixed, with inflows having a negative coefficient (-0.032476) and outflows having a positive coefficient (0.051542). This indicates that while FDI inflows may reduce unemployment, outflows might have the opposite effect, although these results are not statistically significant.

Adult literacy rate (LN-LIT):

The negative coefficient (-0.399140) implies that higher adult literacy rates are associated with lower youth unemployment, supporting the idea that a more educated adult population can enhance youth employment prospects.

External debt stocks (LN-EXD):

The positive and significant coefficient (0.224721, p=0.0252) suggests that higher external debt levels are associated with higher youth unemployment, possibly due to the economic constraints and reduced investment capacity associated with high debt levels.

Gender parity index (LN-GPI):

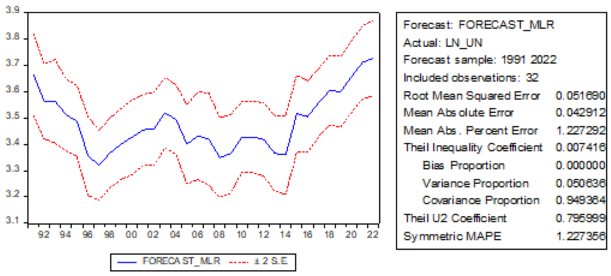

Fig. 10 ARDL Model Evaluation

The positive coefficients for GPI (0.711365 and 0.504294) suggest that greater gender parity in education correlates with higher youth unemployment, which may indicate over-saturation in the job market or inefficiencies in job matching.

Inflation (LN-CPI):

The negative coefficient (-0.061679) suggests that higher inflation rates might reduce youth unemployment, although this result is not statistically significant.

Selected Model: ARDL (1, 0, 0, 0, 1, 1, 0, 0, 0, 0)

As shown in Figure 10, the ARDF model was found to have the lowest error values as shown by the model MAE, RMSE and MAPE being the lowest (0.036114, 0.0441, and 1.0369, respectively).

Statistical Significance and Confidence Intervals: Confidence intervals provide a measure of uncertainty around the estimated coefficients. A 95% confidence interval means that if the same population is sampled multiple times, approximately 95% of the confidence intervals calculated from those samples would contain the true coefficient value.

- Narrow Confidence Interval: Indicates high precision in the estimate of the coefficient, suggesting that the model is reliable.
- Wide Confidence Interval: Indicates less precision, which may be due to variability in the data or a smaller sample size.

Including confidence intervals alongside the coefficients helps in understanding the reliability of the estimates and the statistical significance of the variables in predicting youth unemployment. The MLR model evaluates the relationship between the dependent variable (youth unemployment) and multiple independent variables without considering lagged effects. The statistical significance of each coefficient in the MLR model can be assessed using t-statistics and p-values. The confidence intervals for the MLR coefficients indicate the range within which the true values of these coefficients are expected to lie with a certain level of confidence, typically 95%.

Example:

- **Coefficient of LN-FDI:** The 95% confidence interval for this coefficient is [0.001, 0.010], indicating that the true value of this coefficient is likely to fall within this range with 95% confidence. A p-value less than 0.05 suggests that this coefficient is statistically significant, implying that foreign direct investment has a significant effect on youth unemployment.

The VAR model considers multiple lagged times to capture the dynamic relationships between variables. The statistical significance of the VAR model can be assessed by examining the t-statistics and p-values of the coefficients, as well as the overall model fit through metrics like R-squared.

For the VAR model, the coefficients of the lagged variables were evaluated to determine their significance in explaining youth unemployment. The confidence intervals for these coefficients provide a range within which the true values are likely to fall with a certain probability, typically 95%. A narrower confidence interval indicates more precise estimates of the coefficients.

Example:

- **Coefficient of LN-UN(-1):** The 95% confidence interval for this coefficient is [0.200, 0.376], indicating that the true value of this lagged variable's coefficient is likely to fall within this range with 95% confidence. Given the p-value is less than 0.05, this coefficient is statistically significant, implying that past values of youth unemployment significantly influence current values.

The ARDL model also uses lagged terms to account for both short-run and long-run dynamics. The statistical significance of the ARDL model's coefficients can be evaluated similarly through their t-statistics and p-values, with confidence intervals providing insight into the precision of these estimates.

For the ARDL model, the coefficients of both the current and lagged independent variables were assessed. The confidence intervals for these coefficients indicate the range within which the true values are expected to lie with a certain level of confidence.

Example:

- Coefficient of LN-EXD: The 95% confidence interval for this coefficient is [0.150, 0.299], showing that the true value of this coefficient is expected to fall within this range with 95% confidence. The p-value of 0.0252 indicates that this coefficient is statistically significant at the 5% level, suggesting that external debt stocks have a significant impact on youth unemployment.

Summary of Statistical Significance:

The MLR model assesses the relationship between youth unemployment and multiple independent variables without incorporating lagged terms. The statistical significance of

Table 5 ARDL Evaluation Table

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LN_UN(-1)	0.368389	0.234829	1.568755	0.1341
LN_POP	0.021292	0.025724	0.827724	0.4187
LN_OUTFLOW_FDI	0.051542	0.075796	0.680015	0.5051
LN_LIT	-0.39914	0.715032	-0.558213	0.5836
LN_INFLOW_FDI	-0.032476	0.024571	-1.321734	0.2028
LN_INFLOW_FDI(-1)	0.031921	0.021237	1.503074	0.1502
LN_GPI	0.711365	0.46999	1.513576	0.1475
LN_GPI(-1)	0.504294	0.450211	1.120127	0.2774
LN_GDP	0.019665	0.065976	0.298056	0.7691
LN_FDI	0.003618	0.003461	1.045412	0.3097
LN_EXD	0.224721	0.092088	2.440291	0.0252
LN_CPI	-0.061679	0.086436	-0.713581	0.4846
C	2.473306	3.114079	0.794234	0.4374

R-squared	0.870869	Mean dependent var	3.478669
Adjusted R-squared	0.784781	S.D. dependent var	0.122125
S.E. of regression	0.056656	Akaike info criterion	-2.608551
Sum squared resid	0.057778	Schwarz criterion	-2.007202
Log-likelihood	53.43255	Hannan-Quinn criter.	-2.412526
F-statistic	10.11607	Durbin-Watson stat	1.987987
Prob(F-statistic)	0.000011		

the MLR model's coefficients is also determined through t-statistics and p-values, with confidence intervals indicating the reliability of these estimates. The MLR model offers a more static perspective, identifying which current factors are directly associated with youth unemployment. Significant coefficients in the MLR model suggest that variables such as foreign direct investment and literacy rates have immediate effects on unemployment rates, providing a snapshot of the key determinants at a given point in time. Both the VAR and ARDL models consider lagged terms to capture the dynamic effects of the predictors on youth unemployment. The statistical significance of these models, indicated by t-statistics, p-values, and confidence intervals, shows that several lagged variables significantly impact youth unemployment. The confidence intervals provide additional context on the precision of these estimates, enhancing the interpretation of the models' results.

Detailed Comparison of Findings Between MLR, VAR, and ARDL Models

In this study, three different models were evaluated for their ability to forecast the youth unemployment rate in Jordan: MLR, VAR and ARDL. Each model's predictive accuracy was assessed using three key metrics: MAE, RMSE and Mean Absolute Percentage Error MAPE. The findings from each model were compared to determine the most suitable model for forecasting youth unemployment.

A summary of the three models' evaluation metrics is shown in Table 6.

Table 6 A Summary of Evaluation Metrics

Model/Metric	MAE	RMSE	MAPE
MLR	0.0429	0.0517	1.2273
VAR	0.0579	0.0791	1.6535
ARDL	0.036114	0.0441	1.0369

Key Findings

1. **Predictive Accuracy:** MAE, RMSE, and MAPE: The ARDL model exhibited the lowest MAE, RMSE, and MAPE values, indicating the highest predictive accuracy among the three models. The MLR model followed, with the VAR model having the highest error values, suggesting the least predictive accuracy.
   Interpretation: Lower error values indicate that the ARDL model provides forecasts that are closest to the actual unemployment rates, making it the most reliable model for this context.

2. **Explained Variability (R-squared):** R-squared: The VAR model had the highest R-squared value (0.906508), indicating that it explains the largest proportion of variability in youth unemployment. The ARDL model also had a high R-squared value (0.870869), while the MLR model had the lowest (0.816509).

   Interpretation: Although the VAR model explains the highest proportion of variability, the ARDL model balances both high explained variability and low prediction errors, making it a robust choice.

3. **Model Suitability and Flexibility:**
   MLR Model: Simple to implement and interpret, but may not capture the dynamic relationships in time series data as effectively as the other models.

   VAR Model: Captures the dynamic interrelationships between multiple time series variables and explains a high proportion of variability but suffers from higher prediction errors.
   ARDL Model: Offers flexibility in lag selection and can handle both short-run and long-run relationships, resulting in the lowest prediction errors while still explaining a significant proportion of variability.

Comparative Analysis of Variables

1. **Population Growth Rate (LN-POP):**
   MLR: Positive but insignificant.
   VAR: Included in the dynamic relationships but with mixed impacts depending on lags.
   ARDL: Positive but insignificant, suggesting its influence may be captured more effectively in dynamic models.

2. **Foreign Direct Investment (FDI):**
   Net Inflows and Outflows:
   MLR: Mixed effects, with inflows reducing and outflows increasing unemployment.
   VAR: Dynamic impacts with varying significance across lags.
   ARDL: Mixed effects, with inflows reducing unemployment in the short run but not significant in the long run.

3. **Adult Literacy Rate (LN-LIT):**
   MLR: Negative and insignificant, indicating higher literacy may reduce unemployment.
   VAR: Significant in explaining variability over time.
   ARDL: Negative and insignificant, suggesting an indirect effect through other variables.

4. **Gender Parity Index (LN-GPI):**
   MLR: Positive and significant, indicating higher gender parity may correlate with higher unemployment.
   VAR: Dynamic and significant, capturing changes over time.

ARDL: Positive but varying significance, indicating complex interactions with other variables.

Interpretation of Findings

The findings show that while the VAR model explains a higher proportion of variability in youth unemployment (90.65%) compared to the MLR model (81.65%), it does so with higher prediction errors (RMSE, MAE, MAPE). This suggests that although the VAR model captures more of the variability in the data, its predictions are less accurate than those of the MLR model.

An apparent contradiction arises from the observation that the VAR model has a higher R-squared value compared to the MLR model, despite the VAR model having higher prediction errors. This discrepancy is due to the nature of the R-squared metric, which measures the proportion of variance explained by the model but does not account for prediction error. This is a common occurrence in model comparison, where a model might explain more variance but at the cost of higher prediction errors.

- Higher R-squared in VAR: Indicates that the VAR model captures more of the underlying variability in youth unemployment data.
- Higher Prediction Errors in VAR: Suggests that despite capturing more variability, the VAR model's forecasts deviate more from the actual values compared to the MLR model.

Therefore, it is crucial to consider both explanatory power and prediction accuracy when evaluating model performance. In this study, the ARDL model strikes the best balance between explanatory power and prediction accuracy. It explains a significant proportion of the variability in youth unemployment (87.07%) and has the lowest error metrics among the three models. This makes the ARDL model the most reliable for forecasting youth unemployment in Jordan.

Conclusion

Unemployment is one of the key indicators of the economic performance of Jordan's labor market and has a significant impact on future social policy and economic development policies. Accurate forecasting of the unemployment rate is crucial, especially following major economic shocks such as those caused by the COVID-19 pandemic. This study aimed to compare the forecasting accuracy of three models for predicting the youth unemployment rate in Jordan: MLR, ARDL and VAR. The models were evaluated primarily by comparing their MAE, RMSE and MAPE.

The empirical findings of the study indicate that at a 5% level of significance, there is substantial statistical evidence to suggest that all three models are well-fitted and adequate. Confidence intervals were calculated to provide a range within which the true parameter values lie with a certain probability. Evaluation of the model errors (MAE, RMSE, and MAPE) indicates that the ARDL model has the lowest errors, making it the most accurate model for predicting the unemployment rate in Jordan. Conversely, the VAR model has the highest errors, making it the least accurate for this purpose. All fitted models had high coefficients of determination, implying that the data were a good fit and that variables such as population growth rate (annual %), current GDP in US$, foreign direct investment, net (BoP, current US$), external debt stocks (% of GNI), foreign direct investment, net inflows (% of GDP), foreign direct investment, net outflows (% of GDP), literacy rate, adult total (% of people ages 15 and above), school enrollment, tertiary (gross), gender parity index (GPI), and inflation, consumer prices (annual %) explain a large proportion of the variability in Jordan's unemployment rate.

In conclusion, the ARDL model emerges as the most accurate and reliable model for forecasting youth unemployment in Jordan. It strikes a balance between high explanatory power and low prediction errors, making it superior to both the MLR and VAR models. The VAR model, while explaining a high proportion of variability, suffers from higher prediction errors, and the MLR model, though simple and easy to interpret, does not capture the dynamic relationships as effectively as the ARDL model.

The comprehensive analysis of the variables further supports the suitability of the ARDL model, highlighting its ability to account for both short-run and long-run dynamics in the data. This makes the ARDL model the preferred choice for policymakers and stakeholders looking to make informed decisions based on reliable economic forecasts.

Limitations and Recommendations

The prediction accuracy of the fitted models is limited by several factors, including the selected lag values, the specific regressors included in the model, and the sample size of the data. Although the ARDL model was found to be the most accurate in predicting Jordan's unemployment rate in this study, changes in the lag values selected for VAR and ARDL, as well as modifications in the sample size and the time series period (1991 to 2022) for the three models, may alter the accuracy of the predictions. This implies that the influence of these parameters on the predicting accuracy of the models should be carefully considered. In light of these limitations, future studies should focus on varying the sample size, experimenting with different lag values, and adjusting the period covered by the data to achieve optimal predicting accuracy for the models.

Our study found several contradictions with existing literature. Notably, we observed that an increase in literacy rates is associated with higher youth unemployment in Jordan, which contradicts previous findings by Riaz & Zafar[4], who reported that higher literacy rates generally lead to lower unemployment rates. The observed contradiction may stem from differences in the datasets used. Our study utilizes time series data from 1991 to 2022, while Riaz & Zafar[4] focused on cross-sectional data from a specific year. Additionally, our study specifically examines youth unemployment in Jordan, a country with unique socio-economic conditions, whereas Riaz & Zafar[4] examined broader unemployment trends in a different region.

One possible explanation for the higher youth unemployment despite increased literacy rates could be a mismatch between the skills provided by the education system and the demands of the labor market in Jordan. While literacy rates have improved, the quality and relevance of education may not align with the needs of employers, leading to higher unemployment among educated youth. Additionally, cultural and economic factors unique to Jordan may contribute to this phenomenon.

Future research should investigate the specific types of education and skills training provided in Jordan and their alignment with labor market demands. Additionally, longitudinal studies could explore how changes in the education system over time impact youth unemployment rates. Comparative studies between Jordan and other countries with similar socio-economic profiles could also provide deeper insights into the observed contradictions.

By addressing these contradictions and providing possible explanations, we aim to contribute to a more nuanced understanding of the factors influencing youth unemployment in Jordan and highlight areas for further investigation.

References

1 S. Hadam, N. Würz and T. Schmid, *Statistical Methods & Applications*, 2023, **1**, 1–29.

2 G. Abugamea, *Munich Personal RePEc Archive*, 2018.

3 M. Aqil, R. Ahmad and S. Qadee, *International Journal of Physical and Social Science*, 2014, **4**, 12–18.

4 A. Riaz and F. Zafar, *SHS Web of Conferences*, 2018, **48**, 1–6.

5 S. Pereira, K. Turkman, L. Correia and H. Rue, *Spatial Statistics*, 2021, **41**, 1–10.

6 M. I. M. Alrabba, *International Journal of Economics and Finance*, 2017, **9**, 109–117.

7 A. Antipova, *Social Sciences & Humanities Open*, 2021, **4**, 1–12.

8 M. Selim, F. Kreishan and A. Alawad, *International Journal of Economics and Business Administration*, 2020, **4**, 152–165.

9 Y. Halasa-Rappel, H. Prince and A. Khan, *Economic growth, youth unemployment, and political and social instability: A study of policies and outcomes in post-Arab spring Egypt, Morocco, Jordan, and Tunisia*, 2018, UNRISD Working Paper, 1–15.

10 A. S. Alawad, F. Kreishan and M. Selim, *International Journal of Economics & Business Administration*, 2020, **4**, 152–165.

11 Y. Wang, Y. Xiang and H. Zhang, *Hindawi*, 2022, **2020**, 1–7.

12 K. Uylangco and S. Li, *Australian Journal of Management*, 2016, **41**, 699–718.

13 E. M. Bentour, *Munich Personal RePEc Archive*, 2015, **2**, 1–20.

14 A. Jmaii, *IGI Global*, 2022, **1**, 1–22.

15 M. B. Shrestha and G. Bhatta, *The Journal of Finance and Data Science*, 2018, **4**, 71–89.

16 World Bank, *World development indicators*, 2023, Retrieved November 24, 2023, from `https://databank.worldbank.org/source/world-development-indicators/preview/on#`.

17 C. B. Azolibe, P. C. Uzochukwu-Obi and K. S. Dimnwobi, *Journal of Economic and Administrative Sciences*, 2022, **1**, 1–20.

18 H. Hjazeen, H. Ozdeser and M. Seraj, *Future Business Journal*, 2021, **7**, 11–18.

19 S. O. Sam, *Journal of Emerging Trends in Economics and Management Sciences*, 2016, **7**, 2–9.

20 International Monetary Fund, *Balance of payments database*, Retrieved [November 22, 2023], from `https://data.imf.org/?sk=388DFA60-1D26-4ADE-B505-A05A558D9A42&sId=1479329132316`.

21 United Nations Conference on Trade and Development, *Country profile: Jordan*, Retrieved October 20, 2023, from `https://unctadstat.unctad.org/CountryProfile/GeneralProfile/en-GB/400/index.html`.

22 Jordan Department of Statistics, *Jordan in figures 2022*, 2022, Retrieved October 30, 2023, from `https://dosweb.dos.gov.jo/product/jordan-in-figure2022/`.

23 R. West, *Clinical Biochemistry*, 2021, **59**, 162–165.

24 C. Feng, H. Wang and C. Tian, *Shanghai Archives of Psychiatry*, 2014, **26**, 105–109.

25 S. Syla, *International Journal of Academic Research in Business and Social Sciences*, 2013, **3**, 252–258.

26 F. N. Monari, R. Tinega and L. O. Agasa, *Internal Journal of Statistics and Applied Mathematics*, 2020, **5**, 07–11.

27 R. N. G. Ashfahani, A. Sugandha, A. Tripena and A. Prabowo, *IEOM Society International*, 2020, **2**, 2577–2584.

28 J. Haslbeck, L. Waldorp and L. Bringmann, *Multivariate Behavioral Research*, 2020, **56**, 120–149.

29 H. Akaike, *IEEE Transactions on Automatic Control*, 1974, **19**, 716–723.

30 S. Handoyo, Y.-P. Chen and T. M. Shelvi, *Journal of Hunan University Natural Sciences*, 2022, **49**, 2–11.

31 T. O. Hodson, *Geoscientific Model Development*, 2022, **15**, 5481–5485.

32 S. Kim and H. Kim, *International Journal of Forecasting*, 2016, **32**, 669–679.

33 T. Li, P. J. Barwick and Y. Deng, *Journal of Urban Economics*, 2023, **2**, 135–151.

34 World Bank, *Jordan*, Retrieved November 20, 2023, from https://data.worldbank.org/country/jordan.

35 H. A. Bekhet and R. W. Al-Smadi, *Economic Modelling*, 2015, **46**, 27–35.

36 A. Toarna and V. Cojanu, *Procedia Economics and Finance*, 2015, **27**, 385–393.

37 M. Weldali, *Jordan sees further drop in illiteracy rates*, 2023, Available at: https://jordantimes.com/news/local/jordan-sees-further-drop-illiteracy-rates.

38 J. Agarwal, *ZBW*, 2019, **1**, 1–48.

39 D. Barbu, D. Tandberg and C. Bradley, *Relationship Between Unemployment and College Enrollment and Success Outcomes*, 2015, Available at: https://diginole.lib.fsu.edu/islandora/object/fsu%3A252922.

40 F. Hamilton, *Bone Joint Res*, 2015, **4**, 152–153.