MODELING THE EFFECT OF DEVOLUTION ON YOUTH UNEMPLOYMENT RATES IN KENYA USING AUTOREGRESSIVE INTEGRATED MOVING AVERAGE - INTERVENTION MODEL

Usolo Samuel Wesonga

A thesis Submitted to the School of Natural Sciences in Partial Fulfilment of the Requirements for the Award of Degree of Master of Science in Statistics of Masinde Muliro University of Science and Technology

November, 2025

DECLARATION

This thesis is my original work, prepared with none other than the indicated sources and support, and has not been presented elsewhere for a degree or any other award.

Signature ..... Date .....

Usolo Samuel Wesonga

SES/G/01-70319/2022

CERTIFICATION

The undersigned certify that they have read and hereby recommend for acceptance of Masinde Muliro University of Science and Technology a thesis entitled "Modeling the effect of devolution on youth unemployment rates in Kenya using Autoregressive Integrated Moving Average - Intervention model."

Signature ..... Date .....

Dr. Okoth Annette

Department of Mathematics

Masinde Muliro University of Science and Technology

Signature ..... Date .....

Dr. David Angwenyi

Department of Mathematics

Masinde Muliro University of Science and Technology

COPYRIGHT

This thesis is a copyright material protected under the Berne convention, the copyright Act of 1999 and other international and national enactments in that behalf, on intellectual property. It may not be reproduced by any means in full or in parts except for short extracts in fair dealing so for research or private study, critical scholarly review or disclosure with acknowledgement, with written permission of the Director Postgraduate Studies on behalf of both the author and Masinde Muliro University of Science and Technology.

DEDICATION

To my late dad, Usolo Joseph Wekesa, whose spirit continues to guide me; to my mother, Jane Alois, whose love and prayers have strengthened me; and to all my relatives and friends who have stood by me. Your support and faith in me fueled my determination towards achieving this important milestone that required hard work and sacrifice.

ACKNOWLEDGEMENT

I acknowledge God as the source of all wisdom and strength. He knows all and does all. I express my gratitude to my respected supervisors, Dr. Okoth Annette and Dr. David Angwenyi, for their invaluable insights that added quality to this work and brought it to this level. Their dedication, commitment and trust in me gave me the impetus to always go an extra mile to accomplish the tasks assigned. I am indebted to the entire department of Mathematics (MMUST) for their support throughout the journey.

ABSTRACT

Youth unemployment remains a major concern, particularly in African countries with the youngest population globally. In Kenya, youth unemployent rate has shown fluctuations despite several government efforts such as *Kazi kwa Vijana* (jobs for youth), Youth Enterprise Development Fund (YEDF), Kenya Youth Empowerment Project (KYEP) and Youth Employment Scheme Abroad (YESA). There is limited empirical research that has quantified the effect of devolution on youth unemployment trends in Kenya, as seen in the literature. To address this gap, the study modeled the effect of devolution on youth unemployment rates using an Autoregressive Integrated Moving Average (ARIMA) - Intervention model. ARIMA - Intervention model was selected due to its applicability in quantifying the effect of policy changes on the time series of interest. To capture the time varying effect of devolution on youth unemployment while controlling the effect of key macroeconomic indicators such as Gross Domestic Product, inflation rates, interest income on the Kenya's public debt and Foreign Direct Investment (FDI), the Dynamic Adaptive Intervention Model (DAIM) was used. Secondary data on youth unemployment rates from the World Bank data bank covering the years 1991 to 2022 was used. Computational analysis was done using Python and R programming software, with model selection based on the Akaike Information Criterion (AIC) and Pseudo R - Squared for ARIMA model and DAIM respectively. ARIMA(0,0,0)(0,0,1)[4] model with non-zero mean incorporating devolution as an intervention revealed that the effect of devolution on youth unemployment rate was statistically significant at 5% significance level. The average unemployment rate rose from 6.67% during the pre-intervention period to 10.19% during the devolution period. The counter factual projection was approximately 8.583%, suggesting that devolution had a measurable and negative effect on youth unemployment rates. The fitted DAIM revealed that the negative effect of devolution on youth unemployment intensified gradually over time, rather than manifesting as an immediate shock. Additionally, interest income on the Kenya's public debt consistently worsened the time-varying effect of devolution on youth unemployment rates across all the considered distributional quantiles. Empirically, the study findings provides a quantitative basis for evaluating the effectiveness of the devolved governance system in addressing youth unemployment in Kenya.

TABLE OF CONTENTS

TITLE PAGE i
DECLARATION ii
COPYRIGHT iii
DEDICATION iv
ACKNOWLEDGEMENT v
ABSTRACT vi
LIST OF TABLES x
LIST OF FIGURES xi
LIST OF ABBREVIATIONS AND ACRONYMS xii

CHAPTER ONE : INTRODUCTION **1**
1.1 Background Information 1
1.2 Statement of the Problem 4
1.3 Objectives of the Study 5
1.3.1 Main Objective 5
1.3.2 Specific Objectives 5
1.4 Justification of the Study 5
1.5 Methods 6
1.5.1 Source of Data 6
1.5.2 Assumptions of the Study 6

CHAPTER TWO : LITERATURE REVIEW **7**
2.1 Introduction 7
2.2 Youth Unemployent 7
2.3 Devolution 11
2.4 Theoretical Review 12
2.4.1 Seasonal ARIMA (SARIMA) Model 13
2.4.2 ARIMA – Intervention Model 13
2.4.3 Dynamic Adaptive Intervention Models (DAIM) 15

2.5 Empirical Review 16

CHAPTER THREE : MODEL FITTING **19**

3.1 Introduction 19

3.2 ARIMA - Intervention Model Identification 19

3.2.1 Pre-Intervention (Noise) Model 20

3.2.2 Estimation of ARIMA - Intervention Model Parameters . 24

3.2.3 Forecasting 27

3.3 Dynamic Adaptive Intervention Model (DAIM) 27

CHAPTER FOUR : RESULTS **30**

4.1 Introduction 30

4.2 Descriptive Statistics 30

4.3 Pre-Intervention (Noise) Model 31

4.3.1 Testing for Stationarity 31

4.3.2 Model Identification 35

4.3.3 Model Comparison and Selection 36

4.3.4 Estimated Model Parameters 37

4.3.5 Model Diagnostic Checking 37

4.4 ARIMA-Intervention Model 40

4.4.1 Model Diagnostic Checking for ARIMA - Intervention Model 41

4.5 Forecasted Youth Unemployment Rates (2023–2030) 43

4.6 Comparative Analysis 45

4.6.1 Evaluating the Effect of Devolution on Youth Unemployment Rates Using Counterfactual Trends 46

4.6.2 Youth Unemployment Rates before and during Intervention 47

4.7 Fitted Dynamic Adaptive Intervention Model 49

4.7.1 Testing for Normality Assumption 49

4.7.2 Estimated Dynamic Effect of Devolution on Youth Unemployment Rates 50

4.7.3 Estimated Effect of Macroeconomic Variables 51

4.7.4 Estimated Dynamic Effect of Devolution Based on the Fitted Models 54
4.7.5 Model Selection 56

CHAPTER FIVE : CONCLUSION AND RECOMMENDATIONS 57

5.1 Introduction 57
5.2 Conclusion 57
5.3 Recommendations 58
REFERENCES 59
APPENDICES 63

LIST OF TABLES

Table 4.1: Summary statistics on youth unemployment rates. 30
Table 4.2: Comparison of the selected seasonal ARIMA models based on their AIC values 36
Table 4.3: Parameter estimates of ARIMA(0,0,0)(0,0,1)[4] model. . 37
Table 4.4: Ljung-Box test results for model residuals. 40
Table 4.5: Estimated parameters of ARIMA(0,0,0)(0,0,1)[4] model incorporating devolution intervention. 40
Table 4.6: Ljung - Box test results for model residuals, for white noise assumption. 43
Table 4.7: Quarterly point forecasts for youth unemployment rates with their corresponding 95% confidence intervals. 44
Table 4.8: Estimated dynamic effect of devolution on youth unemployment, $\hat{\beta}_t$ and uncertainty, P_t values before accounting for macroeconomic variables 50
Table 4.9: Quantile regression of the estimated dynamic effect of devolution at 25th percentile level ($\tau = 0.25$) 52
Table 4.10: Quantile regression of the estimated dynamic effect of devolution at 50th percentile level ($\tau = 0.50$) 53
Table 4.11: Quantile regression of the estimated dynamic effect of devolution at 75th percentile level ($\tau = 0.75$) 54
Table 4.12: Estimated dynamic effect of devolution on youth unemployment rates, $\hat{\beta}_t$ at 25th, 50th and 75th percentile levels 55
Table 4.13: Comparison of Pseudo R-Squared for quantile regression at 25th, 50th and 75th percentile levels 56

LIST OF FIGURES

Figure 1.1: Population pyramid of Kenya, 2019. 3

Figure 4.1: Time series plot of observed youth unemployment rates before the devolution period. 32

Figure 4.2: ACF plot of the youth unemployment rates before devolution. 33

Figure 4.3: PACF plot of the youth unemployment rates before devolution. 33

Figure 4.4: Seasonal ACF plot of the observed youth unemployment rates before devolution. 34

Figure 4.5: Seasonal PACF plot of the observed youth unemployment rates before devolution. 35

Figure 4.6: ACF plot of residuals. 38

Figure 4.7: Normal Q - Q plot of residuals. 38

Figure 4.8: Histogram of residuals overlaid with a normal curve. . . 39

Figure 4.9: Histogram of residuals for ARIMA-Intervention model. . 41

Figure 4.10: Q-Q plot of residuals for ARIMA-Intervention model. . 42

Figure 4.11: The observed historical youth unemployment rates and the 8 year forecasts (2023 - 2030) 45

Figure 4.12: Observed and predicted youth unemployment rates with the 2013 intervention point. 46

Figure 4.13: A flow chart showing youth unemployment rates before and during the devolution period. 48

Figure 4.14: Histogram of residuals for each economic variable 49

LIST OF ABBREVIATIONS AND ACRONYMS

ACF	Autocorrelation Function
AIC	Akaike Information Criterion
AR	AutoRegressive
ARIMA	Autoregressive Integrated Moving Average
DAIM	Dynamic Adaptive Intervention Model
FDI	Foreign Direct Investment
GDP	Gross Domestic Product
ILO	International Labour Organization
KNBS	Kenya National Bureau of Statistics
MA	Moving Average
NEET	Not in Education, Employment or Training
NGO	Non Governmental Organization
OECD	Organization for Economic Cooperation and Development
PACF	Partial Autocorrelation Function
SARIMA	Seasonal Autoregressive Integrated Moving Average
SETAR	Self Exciting Threshold Autoregressive
SGED	Skewed Generalized Error Distribution
SPSS	Statistical Package for Social Sciences
UN	United Nations
UNICEF	United Nations International Children's Emergency Fund
VAR	Vector Autoregressive
YEDF	Youth Enterprise Development Fund

CHAPTER ONE

INTRODUCTION

1.1 Background Information

Youth unemployment remains a critical socio-economic concern worldwide, drawing significant attention from international development organizations, policymakers and researchers. United Nations (UN) defines youth as persons aged 15 to 24 years, a standard definition adopted globally to ensure comparability of statistics and policy frameworks [48]. This age group cohort is particularly important since it captures the critical school to work transition period, where youth seek to establish their first stable employment. This transition is often characterized by limited job opportunities, skills mismatch and structural barriers that hamper entry into productive employment [17]. For this reason, the World Bank and ILO collect and report labour market indicators using this age range, such as Youth Unemployment Rate and Youth Not in Education, Employment or Training (NEET), so as to provide consistent international benchmarks [52] [18]. Extended unemployment at this age bracket can have long-term contrary effects, such as lower employability, reduced lifetime income and skills depreciation. Moreover, youth unemployment has broader economic and social implications, such as social unrest, increased dependency ratio and slower economic growth [52].

According to the International Labour Organization (ILO), age 15 is considered to be the lower bound for labor force participation according to the most national and international labour statistics. This makes the age bracket 15–24 both a statistically practical and policy relevant class for analyzing labor market outcomes among the youths [16]. Therefore, understanding unemployment trends in this age group provides valuable insights for designing efficient interventions aimed at promoting youth employment, especially in developing economies like Kenya, where the youth population forms a meaningful proportion of the total labour force.

Kenya has experienced a notable Gross Domestic Product (GDP) growth since 2000, whereby between 2006 and 2013, the Kenya's expanding economy saw a yearly employment growth rate of approximately 45%. However, this increase has not been able to absorb approximately 500,000 to 800,000 Kenyan youths entering the labor market each year. The exact number of those entering the labor market remains unclear, leading to significant variance in employment statistics [11]. While government programs such as Youth Enterprise Development Fund (YEDF) and *Kazi Kwa Vijana* were designed to create employment opportunities, they have faced implementation challenges, limiting their effectiveness [22].

Creative fields such as writing, art, music, drama and fashion are frequently seen as pastimes and many parents in Africa often advise their children against choosing careers in these sectors. This sector faces challenges due to mismanagement, dysfunctional guilds and associations, lack of successful entrepreneurship, along with weak collaboration with government agencies. Funding intended to tackle these problems has faced criticism for fostering dependency instead of promoting entrepreneurship [13].

While some young individuals with limited formal education or work experience struggle to find jobs in the formal sector, they still earn a living by working in the informal sector. Informal job sectors encounter various challenges such as crime, limited access to loans and financial services, poor infrastructure, insufficient training centers and limited opportunities for market access and business growth programs [53].

The Kenya National Bureau of Statistics (KNBS) census reports (2009 and 2019) indicated that Kenya's youth population grew from 11,089,518 in 2009 to 13,177,600 in 2019. Consequently, the proportion of youth in the total population slightly increased from 28.7% to 29% [24].

Figure 1.1 presents a population pyramid illustrating the age and gender distribution of Kenya's population in 2019 (*Source: KNBS 2019 Census Report*).

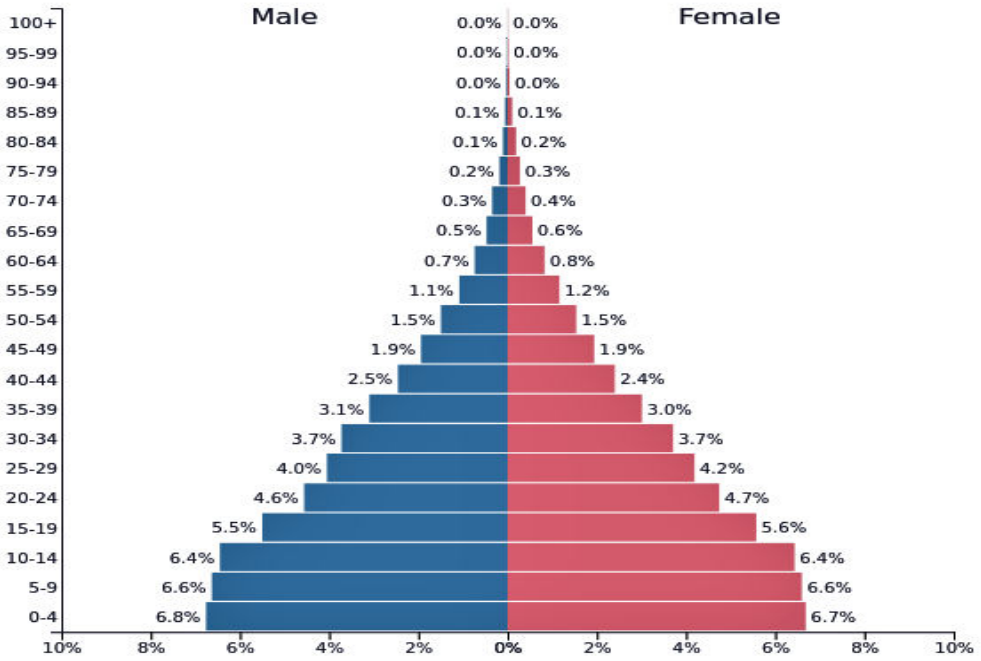

Figure 1.1: Population pyramid of Kenya, 2019.

Figure 1.1 illustrates that young people aged 15 to 34 constitute the largest share of the working-age population, with individuals aged 15 to 24 representing a significant portion.

Introduction of devolution aimed to distribute governance responsibilities to lower administrative levels, ensuring better resource allocation and economic balance. While devolution has achieved some remarkable successes, such as improved infrastructures and localized decision-making, a number of literature has disclosed that its effectiveness in generating job opportunities remains uncertain and its implementation has also faced significant limitations in various sectors [49]. According to United Nations International Children's Emergency Fund (UNICEF), devolution can only yield significant employment opportunities for the youth if the implemented policies prioritize them. Despite increased financial liberty at the county level, youth employment has remained slow, raising concerns about the effectiveness of decentralized governance in addressing youth unemployment [50].

Over the past 30 years, many African countries have worked on devolution reforms, aiming to increase governance efficiency and ensure that policies meets the needs of local communities, especially the marginalized groups. There is however limited studies in Africa explaining how devolution affects service delivery and

reducing poverty. The few pieces of evidence have shown that devolution has not yet achieved the results it promised. A devolved system of governance was introduced in Kenya for the citizens to access public services more easily. It was also viewed as a way of boosting social and economic growth, creating new jobs and investment opportunities that would in turn lead to greater advancement [28] [34].

Kilonzo stated that devolution would make it easier for people to access financial services such as banking, capital markets, fund management and insurance. Also, counties would act as centers for growth, where people can discover chances to invest. These activities would attract both local and international investors, creating more job opportunities for the unemployed youths, hence making it easier for them to access financial services [23]. According to Putnam, devolution can also have unintended outcomes, including governance inefficiencies and increased local-level corruption, which may negatively affect job creation efforts [40].

1.2 Statement of the Problem

Youth unemployment has led to increased poverty, rising crime rates, social exclusion and psychological distress, hindering national development [15]. According to KNBS (2019), youth aged 20 - 24 had the highest unemployment rate (16.8%), followed by those aged 15 - 19 (11.1%). The report further indicated that those aged 20 - 24 had the highest long-term unemployment rate at 8.8%. In 2022, youth aged 20-24 years continued to have the highest long-term unemployment at 9.9% [25] [26].

Research has examined how unemployment affects young people, along with key factors contributing to it. However, little attention has been paid to modeling the effect of devolution on youth unemployment within the Kenyan context using empirical time series modeling. Therefore, this study aimed at modeling the effect of devolution on youth unemployment in Kenya using Autoregressive Integrated Moving Average - Intervention model. ARIMA - Intervention model was selected due it applicability in quantifying the effect of policy changes on the time series of interest. While other factors (e.g. economic recessions) play a role, devolution is a policy issue with an expected employment effect, making it a justified intervention to study.

1.3 Objectives of the Study

1.3.1 Main Objective

To model the effect of devolution on youth unemployment rates in Kenya using Autoregressive Integrated Moving Average - Intervention model.

1.3.2 Specific Objectives

(i) To fit an ARIMA-Intervention model on youth unemployment data in Kenya during devolution.

(ii) To forecast youth unemployment rates with devolution intervention using the fitted ARIMA-Intervention model.

(iii) To quantify the effect of devolution on youth unemployment in Kenya by comparing youth unemployment rates before and during the devolution period.

(iv) To estimate the time-varying effect of devolution on youth unemployment rates while controlling key macroeconomic factors, namely, Gross Domestic Product (GDP), inflation rates, interest income on Kenya's public debt and Foreign Direct Investment (FDI).

1.4 Justification of the Study

Most previous studies on devolution and youth unemployment in Kenya have been largely descriptive or cross-sectional, focusing on institutional and governance aspects rather than on quantifiable economic policy interventions over time. Such approaches fail to adequately capture the dynamics of youth unemployment rates and the delayed effects of policy interventions. A time series approach, particularly one that accounts for intervention effects, provides a more robust framework for assessing the effect of policy reforms. Through this modeling approach, that is, use of ARIMA – Intervention model, determined whether devolution brought about a statistically significant shift in youth unemployment rates.

This study is further justified by its possible contribution to both academic and policy discourse. Empirically, the study provides a quantitative basis for assessing the effectiveness of the devolved governance system in addressing youth unemployment in Kenya. Methodologically, it contributes to the existing literature by applying an advanced time series modeling methods, which is rarely used in policy evaluation within the Kenyan context. The study findings will therefore

enable policymakers, county governments and other development agencies in designing targeted interventions that enhance youth employment creation under the devolved governance model.

1.5 Methods

This study employed quantitative time series modeling techniques to model the effect of devolution on youth unemployment rates in Kenya, that is, using an ARIMA - Intervention model. The methodological framework comprised of three key stages: Data sourcing and preprocessing, model specification, parameter estimation and model diagnostic checking.

1.5.1 Source of Data

The study used secondary data on the yearly youth unemployment rates and the selected macroeconomic indicators (Gross Domestic Product (GDP), inflation rates, interest income on the Kenya's public debt and Foreign Direct Investment (FDI)) obtained from the World Bank databank covering the years 1991 to 2022. The yearly data was then interpolated using a stochastic approach, based on the Skewed Generalized Error Distribution (SGED), which introduced controlled variability and preserved realistic economic patterns. This was done so as to obtain the required sample size for time series modeling. R and python programming software were used for computational analysis.

1.5.2 Assumptions of the Study

(i) The white noise process e_t is normally distributed with mean zero and variance σ_e^2.

(ii) The study assumed a linear relationship between youth unemployment rates and devolution, implying that factors related to devolution policy implementation translates proportionally into changes in youth unemployment rates.

(iii) The random error term η_t in the Dynamic Adaptive Intervention model was assumed to be distribution-free, so as to allow flexibility in capturing random fluctuations without imposing strict normality assumptions.

(iv) This study defined a youth as an individual aged 15 - 24. The definition was adopted based on the availability of data within this age limit.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This chapter reviews the existing literature relevant to this study, aiming to establish the research gap and justify the chosen methodological approach. The structure is as follows: Section 2.2 discusses studies on youth unemployment; Section 2.3 reviews literature related to devolution; Section 2.4 explores the theoretical framework underlying the models employed; and Section 2.5 details the empirical review of the models applied in this study.

2.2 Youth Unemployent

According to the International Labour Organization (ILO) 2024 report, youth consistently face higher unemployment rates than adults, often due to skills mismatch, limited work experience and encounter barriers when trying to enter the labor market. Youth unemployment is even more pronounced in many low and middle income countries, where youth population is growing rapidly with a slow job creation, indicating that youths are competing for the available limited opportunities. For this reason, understanding what drives youth unemployment has become an important area of research and policy discussion, with scholars increasingly examining how both structural conditions and broader economic issues shape youth employment in the labour market [27] [32].

Lam examined youth unemployment trends across 154 countries globally. According to his study findings, although global youth unemployment was approximately at 12.6%, the commonly referenced idea of a "youth bulge" where a large proportion of the population comprises of the youth did not explain why some countries experience higher youth unemployment than others. The study suggests that simply having more young people does not automatically lead to higher unemployment. Instead, it depends much on the conditions that youth face when trying to enter the labor market, such as availability of appropriate skills training, flexibility of labour market institutions and whether economic policies favor job

creation, which implies that demographic trends may set the stage, but it is the labor market structure and policy issues that shape the employment outcomes [27].

Mwakalila used a Generalised Method of Moments (GMM) technique to find out the association between income inequality and youth unemployment from 42 countries in Africa over nearly three decades. He argued that income inequality plays an important role in defining youth unemployment outcomes. The study demonstrated that countries with higher inequality levels tend to have higher youth unemployment, even after accounting for the effect macroeconomic indicators such as foreign direct investments, population growth, economic output and political influence. This implies that besides economic growth, distribution of opportunities and resources matters greatly. Youths in less privileged societies face steeper challenges entering the labour markets, particularly when access to quality education and employment networks remains uneven [32].

The Kenya Institute for Public Policy Research and Analysis (KIPPRA) 2012 report showed that youth, oftentimes defined as individuals aged 15 - 34, continue to face high unemployment compared to older adults. The report highlights that education level, household economic background and working experience all influence the likelihood of finding jobs among the youth [21]. Otoi and Pokhariyal used Autoregressive Distributed Lag (ARDL) modelling approach to further explore these factors the on yearly data from 1979 to 2012. Their results indicated that youth unemployment is influenced in the long run by GDP growth, foreign direct investment and external debt. Their study also revealed an unexpected finding, that economic growth was sometimes associated with an increase in youth unemployment, suggesting that economic growth alone is not sufficient, unless it is accompanied by job creation that matches the skills and needs of the youths. Together, these studies showed that youth unemployment in Kenya is not driven by a single factor, but reflects a complex interplay of economic, social and institutional conditions [39].

Okuom *et. al* used the Autoregressive Distributed Lag (ARDL) mode to analyze the association between Kenya's youth unemployment rate and per capita income was investigated. The model was chosen due to its flexibility in analyzing the short-term and long-term effects of both stationary and non-stationary variables.

The study was anchored on Okun's law, which states that a 2% rise in Gross Domestic Product (GDP) is equivalent to a 1% drop in unemployment (a 1% rise in employment). The World Bank database provided secondary time series data for the study, which covered the years 1991 to 2021. The results showed that youth unemployment rates increased by 4.88% for every 1% increase in per capita income. The surprising conclusion was that Kenya is experiencing *"jobless growth,"* in which economic expansion does not result in a significant number of new jobs [38]. Nevertheless, the effect of devolution on youth unemployment trends using ARIMA - Intervention model was not taken into consideration.

Wamalwa ascertained how non-technological factors, such as labor supply, affect youth employability, a Probit model was used to analyze the determinants impacting youth unemployment in Kenya. The model was used to assess how family background and personal characteristics affect youth employment prospects in Kenya. According to the study findings, there is skills mismatch in the labor market since young people with only secondary education or vocational training had greater unemployment rates compared to those with tertiary education. Key drivers of employment were found to include household factors (financial status and family size) and personal characteristics (age, gender, location and educational attainment) [51]. The study did not, however, look at how devolution has influenced the development of youth employment prospects using a time series methodology.

Sam *et. al* evaluated the macroeconomic variables affecting youth unemployment (YUN) in Kenya using applied cointegration analysis. Gross domestic product (GDP), external debt (ED), private investment (PI), Foreign Direct Investment (FDI), youth population (POP) and youth literacy rate (LR) were considered as the independent variables. According to the study, GDP, external debt, foreign direct investment, private investment, literacy rates and youth population growth all accounted for 84.08% of the differences in youth unemployment. According to the study results, youth unemployment rates rose by 1.3225%, 0.356204%, 0.269%, 0.002441% and 0.154216% respectively, for every 1% increase in GDP, ED, FDI, PI and LR. The *"jobless growth"* theory was further supported by the thought – provoking finding that there was a positive correlation between GDP growth and youth unemployment [44]. The study however, did not quantify the role of devolution in shaping labour market dynamics, particularly for the youth.

Katumo looked at the connection between Kenya's economic growth and youth unemployment. The Granger causality test and the Ordinary Least Squares (OLS) methods were used to demonstrate the relationships between the variables using secondary data. Keynesian theory, Okun's law, the Solow-Swan model and surplus-value theory were among the theories that guided the study. These theories aid in the explanation of economic growth and unemployment. According to the study, Okun's law does not apply to youth unemployment in the Kenyan economic context. The correlation between economic growth and youth unemployment was found to be 0.708, far larger than the 0.03 suggested by Okun's law. Furthermore, Okun's inverse relationship between youth unemployment and economic growth did not hold. The results demonstrated that economic growth by itself is insufficient to explain trends in youth unemployment; thus, it was recommended that additional factors be examined in order to have a more thorough understanding of the changes in youth unemployment [20]. This emphasizes the necessity of investigating policy measures like decentralization and governance in order to alleviate youth unemployment.

Munyao examined the impact of youth unemployment on Africa's economic development. Unlike other purely econometric analyses, Munyao did not use a formal statistical model such as regression or time series analysis, instead, the study relied on content analysis of documentary sources and interpretive reasoning to derive insights. With a sample size of 384 participants, this study used primary data that was collected using questionnaires and interviews. Officials from non-governmental organizations (NGOs) that assist young people in Nairobi, as well as community leaders, youth and religious leaders, were selected through the use of purposeful sampling techniques. Microsoft Excel and the Statistical Package for Social Sciences (SPSS) were used to conduct the analysis. According to the report, growing unemployment impeded economic growth and retarded development. The findings criticized the absence of follow-up procedures in current government programs, while also recognizing Kenya's efforts to combat youth unemployment [31]. The study did not, however, examine how devolution contributes to the development of jobs for youth, using a time series methodology.

Monari *et. al* predicted the Kenyan youth unemployment using a multivariate linear regression model that took into account population growth, GDP and Foreign Direct Investments (FDI). The World Bank databank provided secondary

data for this study, including data on youth population, GDP growth rate and FDI from 2000 to 2017. SPSS software and E-Views 8 were used for data analysis. According to the study findings, youth unemployment and GDP are positively correlated ($B_1 = 0.027$). Additionally, the study discovered that a one-unit shift in FDI would result in a 0.034 unit increase in youth unemployment, provided all other parameters remained constant. On the other hand, youth unemployment would drop by 1.543 units for every unit increase in population growth. According to the study, strategies that draw in foreign direct investment should be taken into account by both national and local governments. These investments can boost GDP growth by adding value and utilizing funds from external debt [30]. However, ARIA - Intervention model was not taken into consideration, using devolution as a policy change in the study.

Riechi investigated the relationship between youth unemployment and insecurity between 2010 and 2018 in Kwale County, Kenya. The research design used in this study was mostly descriptive. Youth leaders, police officers, Kwale County government representatives, Kwale County national government administration personnel and members of the religious community participated in the focus group. Focused group discussions, interviews and questionnaires were employed in the study to gather primary data. According to the survey, majority of the youths in Kwale County were either unemployed or working in low-wage formal occupations like fishing, boat riding, boda riding and small-scale trading [42]. Although this study emphasized the security threats linked to youth unemployment, it did not investigate how devolution has contributed to the issue using a time series methodology, i.e. ARIMA - Intervention model.

2.3 Devolution

The primary justification of devolution is that local governments are better equipped to recognize and respond to the particular requirements of their communities, such as job creation and economic development [43]. Devolution is argued to boost local labor markets in the context of employment by promoting local development projects that generates jobs suited to the unique needs and resources of local communities, fostering participatory governance, which gives citizens a say in policies that impact labor market outcomes and improving resource allocation efficiency, which permits focused investments in sectors like infrastructure, education and employment-generating industries [46] [10] [37].

Nimrod used descriptive research design to assess the effect of devolution on Kenya's economic development, with a particular emphasis on the Nakuru East sub - county. The results of the study demonstrated that the situation had clearly improved following the introduction of devolution. The results of the survey also showed that devolution had improved the social and economic standing of the residents of the Nakuru - East constituency. The study's conclusion demonstrated that devolution greatly enhanced social and economic circumstances, but they also highlighted the necessity of greater policy execution and accountability [35]. However, the study did not quantify the effect of devolution on youth unemployment, using ARIMA - Intervention model.

Nzau determined the impact of decentralized funds on Kenya's economic growth from 1993 to 2012. He used an econometric time series regression model, specifically an Ordinary Least Squares (OLS) framework, to estimate the relationship between devolved funds and economic growth. The study findings supported the arguments made by those who opposed that devolution had led to economic growth. Fiscal decentralization was accomplished through decentralized finance. As recommended by the two distinct schools of thought on fiscal decentralization, it was divided into capital finance and recurrent finance in order to isolate the effects of each on economic growth. All of the model's fiscal decentralization components had a negative correlation with economic growth. This suggested that poor management of devolved resources may have hampered development and that fiscal decentralization had a detrimental effect on economic growth [36]. However, he did not consider ARIMA - Intervention model in achieving his research objectives.

2.4 Theoretical Review

Autoregressive Integrated Moving Average (ARIMA) Model

ARIMA modeling is a helpful tool for assessing the effects of large-scale time series data when other methods fails to work well. It takes into account the underlying trends and autocorrelation, and it offers flexible ways to model various types of impact [33].

ARIMA(p,d,q) is a combination of the Autoregressive(AR) model of order p, moving average model (MA) of order q and the differencing (I) component of order d needed to arrive at stationary $ARMA(p,q)$ model [3].

An $ARIMA(p, d, q)$ model is given by the following expression;

$$(1-B)^d Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - ... - \theta_q e_{t-q} \tag{2.1}$$

Rewriting Equation (2.1) using the backshift operator technique, that is, $Y_{t-k} = B^k Y_t$:

$$(1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p)(1-B)^d Y_t = (1 - \theta_1 B - \theta_2 B^2 - ... - \theta_q B^q) e_t \tag{2.2}$$

$$\phi(B)(1-B)^d Y_t = \theta(B) e_t \tag{2.3}$$

where:

Y_t represents the non stationary time series, $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p$ is the autoregressive component polynomial, $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - ... - \theta_q B^q$ is the moving average component, $(1-B)^d$ is the trend difference operator, B is the backshift operator and e_t is the white noise process.

2.4.1 Seasonal ARIMA (SARIMA) Model

SARIMA model builds on the basic ARIMA model and is specifically created to work with time series data that show both trends and seasonality [14]. This model is useful for the data that experience regular highs and lows at certain times of the year.

$ARIMA(p, d, q)(P, D, Q)[s]$ model is given by;

$$\Phi_P(B^s)\phi(B)(1-B^s)^D(1-B)^d Y_t = \Theta_Q(B^s)\theta(B) e_t \tag{2.4}$$

where;

Y_t represents the non stationary time series, $\phi(B)$ is the usual autoregressive component polynomial, $\theta(B)$ is the moving average component polynomial, $\Phi_P(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - ... - \Phi_P B^{Ps}$ is the seasonal autoregressive component of order P, $\Theta_Q(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - ... - \Theta_Q B^{Qs}$ is the seasonal moving average component of order Q, $(1-B^s)^D$ is the seasonal difference operator, $(1-B)^d$ is the ordinary difference operator and e_t is the white noise process.

2.4.2 ARIMA – Intervention Model

ARIMA - Intervention modeling is a robust statistical technique used to evaluate the effect of a significant policy change or event on a time series of interest. Based on the study by Box and Tiao, it is of interest to determine whether there is any evidence of a change or effect, of an expected kind on the time series associated

with the event [4].

The general structure of ARIMA - Intervention model they developed is of the form;

$$Y_t = \frac{\omega(B)}{\delta(B)} I_{t-b} + \frac{\Theta(B)}{\Phi(B)(1-B)^d} e_t \tag{2.5}$$

Where;

Y_t represents a time series variable of interest (dependent variable). I_t represents the intervention variable. In the case of a step intervention, I_t takes only the values 0 and 1, denoting non-occurrence and occurrence of the intervention event respectively, and it is defined as;

$$I_t = \begin{cases} 1, & \text{if } t \geq T \\ \\ 0, & \text{if } t < T \end{cases} \tag{2.6}$$

And in the case of a ramp intervention;

$$I_t = \begin{cases} f(t-T), & \text{if } t \geq T \\ \\ 0, & \text{if } t < T \end{cases} \tag{2.7}$$

With T being the time of intervention when it first occurred.

$\delta(B) = 1 - \delta_1 B - \delta_2 B^2 - ... - \delta_r B^r$ is the slope parameter. Considering a step intervention, if δ is near 0, the effect of intervention remains constant over time and if is near 1, the effect of intervention increases over time.

$\omega(B) = \omega_0 - \omega_1 B - \omega_2 B^2 - ... - \omega_s B^s$ is the impact parameter, which explains a change due to the intervention effect (the change might be positive or negative).

b is the delay parameter, which usually takes values 0, 1 and so on.

If $b = 0$, it implies that the effect of intervention has occurred at the time of intervention, and if $b = 1$, it implies that the effect of intervention is felt after a delay of one period and so on.

B is the back-shift operator. e_t is the usual white noise process, and they are independently and identically distributed (iid), i.e. $e_t \sim N(0, \sigma^2)$.

For simplicity; we employ the backshift operator to rewrite the intervention component in Equation (2.5), i.e. $I_{t-b} = B^b I_t$, such that;

$$Y_t = \frac{\omega(B)}{\delta(B)} B^b I_t + \frac{\Theta(B)}{\Phi(B)(1-B)^d} e_t \tag{2.8}$$

Further, let $\frac{\omega(B)}{\delta(B)}B^b I_t = \omega$, since it does not contain any variable, i.e. it is a coefficient of the intervention variable I_t; and we know that $\frac{\Theta(B)}{\Phi(B)(1-B)^d}e_t$ follows the Box - Jenkins ARIMA(p,d,q), which forms the noise component of the model. Hence, Equation (2.8) translates to;

$$Y_t = \omega I_t + ARIMA(p, d, q) \tag{2.9}$$

Equation (2.9) forms a simplified ARIMA-Intervention model. This theoretical overview has proven that ARIMA model and its versions, particularly ARIMA Intervention model provide a strong statistical basis for analyzing how interventions such as devolution affect youth unemployment. Even though conventional ARIMA and SARIMA models are accurate at identifying seasonality and temporal trends in unemployment facts, they're not excellent at incorporating structural effects of governance policies. An ARIMA–Intervention framework based on the study of Box and Tiao (1975) gives a methodological advantage by allowing the explicit evaluation of the policy shocks. Therefore, an ARIMA–Intervention model guided this study as the most appropriate theoretical tool to capture the effect of devolution on youth unemployment rates behavior in Kenya.

2.4.3 Dynamic Adaptive Intervention Models (DAIM)

Although some literature has recognized the importance of adaptive or dynamic modeling in behavioral and medical sciences, their application in socio-economic policy evaluation remains limited. Song and Wittes, for instance, emphasized dynamic treatment regimes in longitudinal clinical trials, arguing for time-dependent modeling of effects in adaptive interventions [8] [47]. However, these approaches have not been widely adopted in macroeconomic or labor market studies, where similar dynamics are equally relevant.

The general form of the model is stated as:

$$\beta_t = A\beta_{t-1} + \gamma X + e_t \tag{2.10}$$

where;

β_t denotes the state variable representing the dynamic effect of the intervention at time t.

A is the state transition matrix governing the dependence of the current state on the previous state β_{t-1}, β_{t-1} is the previous dynamic intervention effect, γ is a coefficient vector capturing the influence of the control variables, X is a vector of control variables, and e_t is a noise term representing random fluctuations,

assumed to be white noise with $e_t \sim N(0, \sigma_e^2)$.

2.5 Empirical Review

Karlson and Javed assessed the predictive accuracy of Self-Exciting Threshold AutoRegressive (SETAR), Seasonal Autoregressive Integrated Moving Average (SARIMA) and Vector AutoRegressive (VAR) time series models using unemployment rates data in Sweden. This study aimed at selecting an appropriate and straightforward time series model in producing insightful ideas and recommendations. They used secondary data on unemployment rates from 1983 to 2010 obtained from the Organization for Economic Cooperation and Development (OECD). A recursive approach for out-of-sample assessment used the data from 2011 to 2015 to evaluate the forecasting performance of the models. The study found that, especially in short-term forecasting, the Self-Exciting Threshold AutoRegressive (SETAR) and SARIMA models performed better than the VAR model [19]. Their results indicated that nonlinear and seasonal time series models capture unemployment dynamics more effectively than linear multivariate approaches. However, the study did not incorporate intervention analysis, which could have provided deeper insights into the impact of policy changes on unemployment trends.

Sharahiley conducted a study employing an ARIMA model with three main objectives: to forecast the unemployment rate of Saudi nationals by 2030; to examine the relationship between unemployment, GDP and population growth; and to explain the downward trend in Saudi Arabia's unemployment rate. The dataset used covered annual data on population growth, GDP and unemployment rates from 1999 to 2019. The analysis revealed a significant pattern in the data, making it suitable for forecasting the next decade using an ARIMA(1,0,0) model. The results projected that the unemployment rate would exceed the Saudi Vision 2030 target, reaching 10.99% by 2030. Contrary to conventional economic expectations, both GDP growth and population expansion showed minimal influence on unemployment [45]. However, the study did not quantify the effect of policy factors such as devolution, which may also affect unemployment trends.

Didiharyono and Syukri conducted applied research using quantitative secondary data on unemployment rates from South Sulawesi's Central Statistics Agency. The study objective was to use an ARIMA model to forecast South Sulawesi's

open unemployment rates. Based on their research findings, they concluded that ARIMA (1, 2, 1) was the best time series model for forecasting the future unemployment rates. The results of their study showed that unemployment rates had shown an increasing trend annually. They hoped that the government, especially the South Sulawesi provincial administration, would develop efficient plans and policies to address the unemployment problem. The results underlined the necessity of government initiatives to generate employment opportunities, including trainings in entrepreneurship, the growth of tourism and greater funding for the agricultural and marine sectors [9].

Adenomon modeled Nigerian unemployment rates using ARIMA model. The analysis used annual secondary data from 1972 to 2014. In order to analyze Nigeria's unemployment rates, four distinct ARIMA models were identified and examined. The findings showed that an ARIMA (2, 1, 2) model was the most effective option among the identified options for analyzing Nigeria's unemployment rates over the studied period. The chosen ARIMA model was used to forecast Nigeria's unemployment rates from 2015 to 2018. From 2015 to 2017, Nigeria's projected unemployment rates increased; however, in 2018, they somewhat decreased. The rates stayed rather high between 2015 and 2018. The selected ARIMA model was shown to be stable upon analysis of its diagnostics [2].

Adesegun *et. al* examined the pattern and forecasted Nigeria's unemployment rates using ARIMA model. The primary statistical technique employed was a univariate time series modeling (Autoregressive Integrated Moving Average model). The secondary data obtained from the World Economic Outlook Database, CBN publications, the International Monetary Fund and the NBS bulletin was used in this study. Nigerian unemployment rates from 2000 to 2018 were included in the data. Differentiating was used to achieve stability since the time series plots showed that the unemployment rates varied over time. ARIMA (1,1,1), ARIMA (2,1,1), ARIMA (2,1,2), ARIMA (2,2,1) and ARIMA (2,2,2) were among the various ARIMA models that were taken into consideration. The models had the following log-likelihood and AIC values: (111.08, -50.54,) (109.22, -50.61), (111.36, -50.68) and (107.52, -50.76), respectively. An ARIMA (1, 1, 1) model was selected since it had the highest log likelihood and the lowest AIC value among the examined models. According to the study's findings, the fitted model can be used to forecast, monitor and control Nigeria's unemployment rates. To

combat unemployment, the report suggested curriculum improvements, economic diversification and industrialization [1].

From the reviewed literature, various studies examined the relationship between devolution, youth unemployment and economic indicators, most of them employed descriptive analyses, comparative case studies or regression based approaches. However, none of them has modeled the effect of devolution on youth unemployment rates using time series techniques that account for the effect of policy shocks or structural changes, specifically, the application of an ARIMA - Intervention model which is well suited for quantifying the effect of policy implementation on a time series of interest, has not been explored in this context. This indicates a methodological gap that the present study sought to address by modeling the effect of devolution on youth unemployment rates in Kenya using ARIMA – Intervention model.

CHAPTER THREE

MODEL FITTING

3.1 Introduction

This chapter presents the research methods employed to achieve the study objectives. It describes the procedures used for model identification, estimation, and diagnostic evaluation. The chapter is organized as follows: Section 3.2 details the identification and estimation of the ARIMA–Intervention model; and Section 3.3 discusses the Dynamic Adaptive Intervention Modeling (DAIM) framework.

3.2 ARIMA - Intervention Model Identification

This section outlines the process used in identifying an appropriate ARIMA intervention model to quantify the effect of devolution on youth unemployment rates in Kenya. The observed time series data on youth unemployment rates (1991 - 2022) was modeled as;

$$Y_t = \mu + \omega I_t + ARIMA(p,d,q)(P,D,Q)[4] \tag{3.1}$$

where:

Y_t represents the observed youth unemployment rates, μ is the intercept value, ω is the impact parameter (measures the effect of devolution on youth unemployment rates), I_t is the intervention variable (represents the occurrence of devolution) and $ARIMA(p,d,q)(P,D,Q)[4]$ forms the pre-intervention (noise) model.

It was assumed that the intervention effect of devolution on youth unemployment rates was gradual rather than a sudden structural break, since the effect may take long to manifest. Therefore, the intervention variable I_t was modeled as a ramp function rather than a step function, defined as:

$$I_t = \begin{cases} t - 2013, & \text{if } t \geq 2013 \\ \\ 0, & \text{if } t < 2013 \end{cases} \tag{3.2}$$

3.2.1 Pre-Intervention (Noise) Model

The data on youth unemployment rates before devolution was used to identify a well-specified time series model that captures the noise component. The analysis followed the Box-Jenkins methodology to identify the parameters of the baseline ARIMA(p,d,q)(P,D,Q)[4] model, which served as the foundation for incorporating the effect of intervention in the ARIMA - Intervention model.

The Box-Jenkins methodology included three systematic approaches, that is, model identification, model comparison and selection, and diagnosing to validate the model [5].

a). Model Identification Stage

This stage involved analysing the data on youth unemployment rates before devolution (1991 - 2022) so as to identify the order of autoregressive, moving average and differencing components that best fits the data. It entailed:

1). Testing for Stationarity

Stationarity is a fundamental assumption in time series analysis, ensuring that the statistical properties of a series such as mean and variance remain constant over time. A stationary time series simplifies modeling by maintaining consistency in its behavior. Augmented Dickey-Fuller (ADF) test was employed to determine whether the time series data was stationary. The test detected for the presence of a unit root (a sign of non - stationarity) in the youth unemployment rates, that is, when the statistical properties such as the mean and variance change over time.

The following hypotheses were tested at 5% significance level:

H_0: The series is non-stationary Vs **H_a: The series is stationary.**

2). Exploratory Analysis

A time series plot was used in observing patterns, trends and seasonality in the observed time series data on youth unemployment rates in Kenya before devolution so as to identify whether differencing was needed in achieving stationarity.

3). Autocorrelation Function (ACF)

ACF plot was used to identify the order of a Moving Average (MA) component (q) and in diagnosing the selected model. The plot was obtained by plotting the sample ACF against the time lag k.

4). Partial Autocorrelation Function (PACF)

PACF plot was useful in determining the order of an Autoregressive (AR) component in the model. The order of AR (p) was obtained by examining the points on the plot where the partial autocorrelations for higher lags are essentially zero.

b). Model Comparison and Selection

Several ARIMA models were identified and compared based on their Akaike Information Criterion (AIC) values. The model with the lower AIC was selected as the best-fitting model. This approach ensured a balance between model accuracy and parsimony. AIC is straightforward in that it provides a quick assessment of the model without requiring data splitting, making it more flexible for practical application. Also, it prioritizes models that offer the best predicative accuracy, more lenient in penalizing complexity and often used in situations where the goal is prediction, not necessarily finding the fit model [7].

c). Parameter Estimation

Parameters of the selected model were estimated using the Maximum Likelihood Estimation (MLE) method under the assumption that $e_t \sim (0, \sigma_e^2)$, and they are independently and identically distributed (iid). This involved determining the values of the model parameters that best fit the observed data on youth unemployent rates before devolution. The justification for using MLE estimators is that they are consistent and they have a small mean squared error relative to other competing estimators [6].

Considering the selected ARIMA(0,0,0)(0,0,1)[4] model in Table 4.2, the parameters were estimated as:

$$Y_t = \mu + \Theta e_{t-4} + e_t, \quad \text{where} \quad \mu \in \mathbb{R} \tag{3.3}$$

The expected value and variance of Y_t in Equation (3.3) based on the assumption $e_t \sim N(0, \sigma_e^2)$:

i). Expected value:

$$E(Y_t) = E(\mu) + \Theta E(e_{t-4}) + E(e_t) = \mu \tag{3.4}$$

ii). Variance:

$$Var(Y_t) = Var(e_t) + \Theta^2 Var(e_{t-4}) = \sigma_e^2(1 + \Theta^2) \tag{3.5}$$

Since Y_t is a function of $e_t's$ which are normally distributed, then from Equation (3.4) and (3.5), it was assumed that Y_t is normally distributed, that is, $Y_t \sim$

$N(\mu, \sigma_e^2(1+\Theta^2))$, such that:

$$f(y_t|\mu, \Theta, \sigma_e^2) = \left(\frac{1}{2\pi\sigma_e^2(1+\Theta^2)}\right)^{\frac{1}{2}} exp\left(-\frac{1}{2\sigma_e^2(1+\Theta^2)}(y_t-\mu)^2\right) \tag{3.6}$$

Considering the observed data on youth unemployment rates, $y_t = (y_1, y_2, ..., y_n)$, the likelihood function was defined as;

$$L(\mu, \Theta, \sigma_e^2|y_1, y_2, ..., y_n) = \prod_{t=1}^{n} f(y_t|\mu, \Theta, \sigma_e^2) \tag{3.7}$$

Substituting Equation (3.6) into Equation (3.7);

$$L(\mu, \Theta, \sigma_e^2|y_t) = \prod_{t=1}^{n}\left(\frac{1}{2\pi\sigma_e^2(1+\Theta^2)}\right)^{\frac{1}{2}} exp\left(-\frac{1}{2\sigma_e^2(1+\Theta^2)}(y_t-\mu)^2\right) \tag{3.8}$$

where n is the sample size of the observed time series data on youth unemployment rates;

Solving Equation (3.8), we have;

$$L(\mu, \Theta, \sigma_e^2|y_t) = \left(\frac{1}{2\pi\sigma_e^2(1+\Theta^2)}\right)^{\frac{n}{2}} exp\left(-\frac{1}{2\sigma_e^2(1+\Theta^2)}\sum_{t=1}^{n}(y_t-\mu)^2\right) \tag{3.9}$$

Maximum likelihood estimation seeks μ, Θ and σ_e^2 such that the log-likelihood function is maximal. Considering Equation (3.9), the log-likelihood function is expressed as;

$$l(\mu, \Theta, \sigma_e^2|y_t) = -\frac{n}{2}log(2\pi) - \frac{n}{2}log\sigma_e^2 - \frac{n}{2}log(1+\Theta^2) - \frac{1}{2\sigma_e^2(1+\Theta^2)}\sum_{t=1}^{n}(y_t-\mu)^2 \tag{3.10}$$

To obtain the ML estimate of μ, we partially differentiated $l(\mu, \Theta, \sigma_e^2|y)$ in Equation (3.10) with respect to μ, set the derivative to zero and solved for μ, thus;

$$\frac{\partial}{\partial\mu}l(\mu, \Theta, \sigma_e^2|y_t) = \frac{1}{\sigma_e^2(1+\Theta^2)}\sum_{t=1}^{n}(y_t-\mu) = 0 \tag{3.11}$$

We obtained the MLE of μ as;

$$\hat{\mu} = \frac{1}{n}\sum_{t=1}^{n} y_t \tag{3.12}$$

The ML estimate of σ_e^2 was obtained by partially differentiating Equation (3.10) with respect to σ_e^2, setting the derivative to zero and then solved for σ_e^2;

$$\frac{\partial}{\partial\sigma_e^2}l(\mu, \Theta, \sigma_e^2|y_t) = -\frac{n}{2\sigma_e^2} + \frac{1}{2\sigma_e^4(1+\Theta^2)}\sum_{t=1}^{n}(y_t-\hat{\mu})^2 = 0 \tag{3.13}$$

Solving for σ_e^2;

$$\hat{\sigma_e}^2 = \frac{1}{n(1+\Theta^2)}\sum_{t=1}^{n}(y_t-\hat{\mu})^2 \tag{3.14}$$

To obtain the ML estimate of Θ, Equation (3.10) was partially differentiated with respect to Θ and solved for Θ by equating the derivative to 0:

$$\frac{\partial}{\partial\Theta}l(\mu,\Theta,\sigma_e^2|y_t) = -\frac{n\Theta}{1+\Theta^2} + \frac{\Theta}{\sigma_e^2(1+\Theta^2)^2}\sum_{t=1}^{n}(y_t-\hat{\mu})^2 = 0 \tag{3.15}$$

Solving for Θ;

$$\Theta^2 = \frac{1}{n}\left(\frac{1}{\sigma_e^2}\sum_{t=1}^{n}(y_t-\hat{\mu})^2\right) - 1 \tag{3.16}$$

$$\hat{\Theta} = \sqrt{\frac{1}{n}\left(\frac{1}{\sigma_e^2}\sum_{t=1}^{n}(y_t-\hat{\mu})^2\right) - 1} \tag{3.17}$$

Now, the fitted ARIMA(0,0,0)(0,0,1)[4] model with non zero mean based on the estimated parameters is given by:

$$Y_t = \hat{\mu} + \hat{\Theta}e_{t-4} + e_t \tag{3.18}$$

where;

Y_t represents youth unemployment rate at time $t < 2013$

$\hat{\mu}$ is the estimated intercept value

$\hat{\Theta}$ is the estimated seasonal moving average (MA(1)) parameter

e_t is the white noise process.

d). Diagnostic Checking Stage

Having selected and estimated model parameters, the model residuals were examined to assess the adequacy and stability of the fitted model. This involved checking whether the residuals are white noise, that is, they don't show any patterns (if they are independently and identically distributed) and normally distributed with mean approximately zero.

The Ljung-Box test was applied on the residuals of the selected model under the following hypotheses, at $\alpha = 0.05$:

H_0: The residuals are white noise Vs H_a: The residuals are not white noise.

The Ljung - Box test statistic is computed as [29]:

$$Q^* = n(n+2)\sum_{k=0}^{m}\frac{r_k^2}{n-k} \tag{3.19}$$

where;

n is the number of observations used to estimate the model, m is the maximum

number of lags and r_k is the autocorrelation function of residuals at lag k. The statistic Q^* approximately follows a chi-squared distribution with $n-k$ degrees of freedom, where k is the total number of parameters estimated in the model.

3.2.2 Estimation of ARIMA - Intervention Model Parameters

The parameters of the full ARIMA - Intervention model were estimated over the entire data period (1991 - 2022) using maximum likelihood estimation (MLE) method, under the assumption that $e_t \sim N(0, \sigma_e^2)$. Having identified and selected the suitable noise model at the pre-intervention analysis stage (Equation (3.18)), the regression form of the ARIMA - Intervention model which incorporated devolution policy (I_t) as an intervention was identified as:

$$Y_t = \mu + \omega I_t + \Theta e_{t-4} + e_t \tag{3.20}$$

The parameter estimates for ARIMA(0,0,0)(0,0,1)[4] model incorporating devolution as an intervention variable were obtained using the Maximum Likelihood Estimation (MLE) method.

Expected value of Y_t:

$$E(Y_t) = \mu + \omega I_t \tag{3.21}$$

Variance of Y_t:

$$Var(Y_t) = Var(e_t) + \Theta^2 Var(e_{t-4}) = \sigma_e^2(1+\Theta^2) \tag{3.22}$$

From Equations (3.21) and (3.22), we assumed that Y_t is normally distributed, that is:

$$Y_t \sim N\left((\mu + \omega I_t), \quad \sigma_e^2(1+\Theta^2)\right) \tag{3.23}$$

with probability density function given as;

$$f(y_t|\mu, \omega, \Theta, \sigma_e^2) = \left(\frac{1}{2\pi\sigma_e^2(1+\Theta^2)}\right)^{\frac{1}{2}} exp\left(-\frac{1}{2\sigma_e^2(1+\Theta^2)}(y_t - (\mu + \omega I_t))^2\right) \tag{3.24}$$

We obtained the likelihood function for the observed data, $y_t = (y_1, y_2, ..., y_n)$ as;

$$L(\mu, \omega, \Theta, \sigma_e^2|y_t) = \prod_{t=1}^{n} f(y_1, y_2, ..., y_n|\mu, \omega, \Theta, \sigma_e^2) \tag{3.25}$$

Substituting Equation (3.24) into Equation (3.25):

$$L(\mu, \omega, \Theta, \sigma_e^2|y_t) = \prod_{t=1}^{n} \left(\frac{1}{2\pi\sigma_e^2(1+\Theta^2)} \right)^{\frac{1}{2}} exp\left(-\frac{1}{2\sigma_e^2(1+\Theta^2)} (y_t - (\mu + \omega I_t))^2 \right) \tag{3.26}$$

Solving Equation (3.26), we have:

$$= \left(\frac{1}{2\pi\sigma_e^2(1+\Theta^2)} \right)^{\frac{n}{2}} exp\left(-\frac{1}{2\sigma_e^2(1+\Theta^2)} \sum_{t=1}^{n} (y_t - (\mu + \omega I_t))^2 \right) \tag{3.27}$$

Maximum likelihood estimation seeks μ, ω, Θ and σ_e^2 such that the log-likelihood function is maximal. Considering Equation (3.27), the log-likelihood function was computed as;

$$l(\mu, \omega, \Theta, \sigma_e^2|y_t) = -\frac{n}{2} log(2\pi) - \frac{n}{2} log\sigma_e^2 - \frac{n}{2} log(1+\Theta^2) - \frac{1}{2\sigma_e^2(1+\Theta^2)} \sum_{t=1}^{n} (y_t - (\mu + \omega I_t))^2 \tag{3.28}$$

To obtain the ML estimator of μ, ω, σ_e^2 and Θ, we partially differentiated Equation (3.28) with respect to μ, ω, σ_e^2 and Θ, equated it to 0 and solved for the estimates such that;

MLE for μ;

$$\hat{\mu} = \frac{1}{n} \sum_{t=1}^{n} (y_t - \omega I_t) \tag{3.29}$$

ML estimator of ω was obtained as;

$$\hat{\omega} = \frac{\sum_{t=1}^{n} (y_t - \mu) I_t}{\sum_{t=1}^{n} I_t^2} \tag{3.30}$$

Solving for σ_e^2, we obtained;

$$\hat{\sigma_e}^2 = \frac{1}{n(1+\Theta^2)} \sum_{t=1}^{n} (y_t - (\mu + \omega I_t))^2 \tag{3.31}$$

Solving for the MLE of Θ;

$$\hat{\Theta} = \sqrt{\frac{1}{n} \left(\frac{1}{\sigma_e{}^2} \sum_{t=1}^{n} (y_t - (\mu + \omega I_t))^2 \right) - 1} \tag{3.32}$$

It is complex to solve the Equations (3.29), (3.30), (3.31) and (3.32), since the expressions are functions of other unknown parameters to be estimated. Therefore, the parameter estimates $\hat{\mu}$, $\hat{\omega}$, $\hat{\sigma}_e^2$ and $\hat{\Theta}$ were obtained iteratively by initializing

the parameter estimates in two steps;

Step 1

Initial guess of $\hat{\mu}$ was computed from the data:

$$\mu^{(0)} = \frac{1}{n}\sum_{t=1}^{n} y_t \tag{3.33}$$

and;

$$\hat{\omega} = \omega^{(0)} \tag{3.34}$$

where, the initial guess of $\omega^{(0)}$ was arbitrary chosen to be 0.5.

First iteration:

Substituting the initial guesses $\mu^{(0)}$ and $\omega^{(0)}$ into Equations (3.29) and (3.30), we obtained:

$$\mu^{(1)} = \frac{1}{n}\sum_{t=1}^{n}(y_t - \omega^{(0)} I_t) \quad and \quad \omega^{(1)} = \frac{\sum_{t=1}^{n}(y_t - \mu^{(0)})I_t}{\sum_{t=1}^{n} I_t^2} \tag{3.35}$$

Iteration was done k times until convergence was achieved, that is;

$$|\mu^{(k)} - \mu^{(k-1)}| \approx 0 \quad and \quad |\omega^{(k)} - \omega^{(k-1)}| \approx 0, \quad where \quad k = 1, 2, 3, ...$$

At the kth iteration:

$$\mu^{(k)} = \frac{1}{n}\sum_{t=1}^{n}(y_t - \omega^{(k-1)} I_t) \quad and \quad \omega^{(k)} = \frac{\sum_{t=1}^{n}(y_t - \mu^{(k-1)})I_t}{\sum_{t=1}^{n} I_t^2} \tag{3.36}$$

Step 2

Having obtained the estimates $\mu^{(k)}$ and $\omega^{(k)}$ in Step 1, iteration of $\hat{\Theta}$ and ${\hat{\sigma_e}}^2$ using Equations (3.31) and (3.32) was done k times until convergence was achieved, such that;

$$|\Theta^{(k)} - \Theta^{(k-1)}| \approx 0 \quad and \quad |\sigma_e^{2(k)} - \sigma_e^{2(k-1)}| \approx 0, \quad where \quad k = 1, 2, 3, ...$$

At the kth iteration:

$$\hat{\sigma}_e^{2(k)} = \frac{1}{n(1 + \Theta^{2(k-1)})}\sum_{t=1}^{n}(y_t - (\hat{\mu} + \omega^{(k-1)} I_t))^2 \tag{3.37}$$

$$\hat{\Theta}^{(k)} = \sqrt{\frac{1}{n}\left(\frac{1}{\sigma_e^{2(k-1)}}\sum_{t=1}^{n}(y_t - (\hat{\mu} + \omega^{(k-1)} I_t))^2\right) - 1} \tag{3.38}$$

Having estimated the model parameters in Equations (3.36) and (3.38), the ARIMA-Intervention model for this study was fitted as follows:

$$Y_t = \mu^{(k)} + \omega^{(k)} I_t + \Theta^{(k)} e_{t-4} + e_t \tag{3.39}$$

where;

Y_t represents youth unemployment rate at time t.

$\mu^{(k)}$ is the estimated intercept value

$\omega^{(k)}$ is the estimated impact parameter at the kth iteration

$\Theta^{(k)}$ is the estimated seasonal moving average (MA(1)) parameter at the kth iteration

I_t is the intervention variable which represents the effect of devolution at time t

e_t is the usual white noise process.

3.2.3 Forecasting

Mean Absolute Percentage Error (MAPE) was used to measure the prediction accuracy using the fitted ARIMA - Intervention model. It provided a percentage measure of how far the predicted youth unemployment rates deviate from the observed values.

It is calculated as:

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{Y_t - F_t}{Y_t} \right| * 100 \tag{3.40}$$

where:

Y_t is the actual value at time t, F_t is the forecasted value at time t, and n is the umber of observations

Interpretation:

On average a MAPE of 10% indicates that the predicted values differ from the observed values by 10%, reflecting a reasonably accurate model performance in econometric modeling.

3.3 Dynamic Adaptive Intervention Model (DAIM)

To estimate the time-varying effect of devolution on youth unemployment rates while controlling the effect of key macroeconomic indicators, DAIM was employed. This model allowed the effect of devolution to evolve gradually over time, which is consistent with policy effects that manifest progressively.

State Equation

The dynamic effect of devolution on youth unemployment rates was represented by the parameter β_t, whose formulation was specified as:

$$\beta_t = A\beta_{t-1} + \gamma X_t + \eta_t \tag{3.41}$$

where β_t denotes the effect of devolution on youth unemployment rate at time t, A is the state transition matrix, X_t is a vector of control variables (GDP, inflation rate, interest income on Kenya's public debt and FDI), γ is a vector of parameters associated with the controls and η_t is the state error term.

In this study, the evolution of β_t follows a random walk process, implying that the effect accumulates gradually:

$$\beta_t = \beta_{t-1} + \eta_t \tag{3.42}$$

Observation Equation

The observed youth unemployment rate y_t is linked to the latent state variable β_t by:

$$y_t = D_t\beta_t + \kappa_t \tag{3.43}$$

where D_t is the devolution indicator defined as:

$$D_t = \begin{cases} t - 2013, & \text{if } t \geq 2013 \\ \\ 0, & \text{if } t < 2013 \end{cases} \tag{3.44}$$

and κ_t is the observation noise assumed to be normally distributed.

Kalman Filter Estimation

Estimation of β_t is conducted recursively through the Kalman filter, which consists of a prediction step and an update step.

Prediction Step

We started with initial condition β_t and P_t (at time $t = 0$) as follows:

$$\beta_0 = 0 \tag{3.45}$$

Reflecting no prior dynamic effect of devolution on youth unemployment rates.

$$P_0 = 1 \tag{3.46}$$

Reflecting moderate initial uncertainty.

Using the initial conditions, the next dynamic effect of devolution at time t was predicted using the equation:

$$\beta_{t|t-1} = \beta_{t-1} \tag{3.47}$$

The next uncertainty was predicted using the expression:

$$P_{t|t-1} = P_{t-1} + \sigma_{\xi}^2 \tag{3.48}$$

where σ_{ξ}^2 is the variance of the state disturbance term η_t.

Update Step

After observing the data on youth unemployment rates, we updated the dynamic effect estimates using the weighting factor (Kalman gain) K_t. The Kalman gain K_t was computed as:

$$K_t = \frac{P_{t|t-1} D_t^T}{D_t P_{t|t-1} D_t^T + \lambda^2} \tag{3.49}$$

The updated estimate of the dynamic effect is then:

$$\beta_t = \beta_{t|t-1} + K_t(y_t - D_t \beta_{t|t-1}) \tag{3.50}$$

and the updated uncertainty becomes:

$$P_t = (1 - K_t D_t) P_{t|t-1} \tag{3.51}$$

Interpretation

An increasing β_t indicates that devolution is associated with rising youth unemployment, while a decreasing β_t would imply a positive employment effect. A decreasing P_t reflects growing confidence in the estimated intervention effect as more data accumulate.

CHAPTER FOUR

RESULTS

4.1 Introduction

This chapter presents the empirical results obtained from the application of the ARIMA–Intervention and Dynamic Adaptive Intervention Modeling (DAIM) approaches to quantify the effect of devolution on youth unemployment rates in Kenya. It begins with a summary of the descriptive statistics of youth unemployment rates before and during the devolution period, followed by identification and estimation of the pre–intervention ARIMA model. The subsequent sections detail the fitted ARIMA–Intervention model, forecasted youth unemployment rates for the period 2023–2030, and a comparative analysis of youth unemployment rates before and during the devolution periods. The final sections present the fitted DAIM results.

4.2 Descriptive Statistics

Table 4.1 displays summary statistics of youth unemployment rates before devolution (1991–2012), during devolution (2013–2022) and across the entire period (1991–2022).

Table 4.1: Summary statistics on youth unemployment rates.

Value/Year	**1991-2012**	**2013-2022**	**1991-2022**
Count	88.000	40.000	128.00
Min	4.787	6.144	4.787
Max	9.572	16.27	16.27
Mean	6.673	10.19	7.777
std	0.897	3.120	2.495
25 % (Q_1)	6.145	6.817	6.248
50 % (μ)	6.404	9.999	6.663
75 % (Q_3)	6.943	13.19	8.204

From Table 4.1, before devolution, the rates demonstrated comparably low variability, with a mean of approximately 6.67% and a standard deviation of 0.897. The minimum and maximum rates were approximately 4.79% and 9.57% respec-

tively, suggesting that the rate was relatively stable during this period, showing mild fluctuations around the mean. A smaller inter-quartile range (from 6.145% to 6.943%) further substantiates this stability, indicating that the rates were clustered closely around the median value of 6.40%.

On the other hand, during the devolution period, youth unemployment rates showed a marked increase in both levels and variability. The average rate rose sharply to approximately 10.19%, while the standard deviation increased to 3.12, indicating greater fluctuations in youth employment rates during this period. The maximum rate of 16.27% is nearly double that of the pre-devolution period, highlighting the challenges faced by the youths in labor market during the times of devolved governance. The wide range between the lower quartile (6.817%) and the upper quartile (13.19%) reflects this increased dispersion.

During the entire study period (1991–2022), the average rate stood at approximately 7.777%, with a standard deviation of about 2.495. These aggregated figures masks the pronounced structural shift between the two periods, suggesting that the implementation of devolution tally with an increased variability in youth unemployment rates.

4.3 Pre-Intervention (Noise) Model

This section presents the results of modeling data on youth unemployment rates before the devolution period in order to construct a suitable noise model. The key steps involved were as follows:

4.3.1 Testing for Stationarity

The Augmented Dickey-Fuller (ADF) test was conducted to assess the stationarity of youth unemployment rates data, that is, if the statistical properties such as the mean and variance changed over time.

The following hypotheses at 5% level of significance guided the test:

$$\boldsymbol{H_0}\textbf{: The time series is non - stationary}$$

$$\textbf{Vs}$$

$$\boldsymbol{H_a}\textbf{: The time series is stationary}$$

ADF results:

ADF Stat	Critical value	P-value
-9.0931	-2.850	3.7782e-15

Since the p-value of the ADF test, 3.7782e-15 $\approx$ 0.0000 is less than the significance level of 5%, we therefore reject the null hypothesis. Since the series is already stationary, further transformations like differencing were unnecessary.

b). Time Series Plot

The data on youth unemployment rates prior to devolution were plotted to identify potential trend and seasonality components. Figure 4.1 presents a time series plot for the data on youth unemployment rates before devolution.

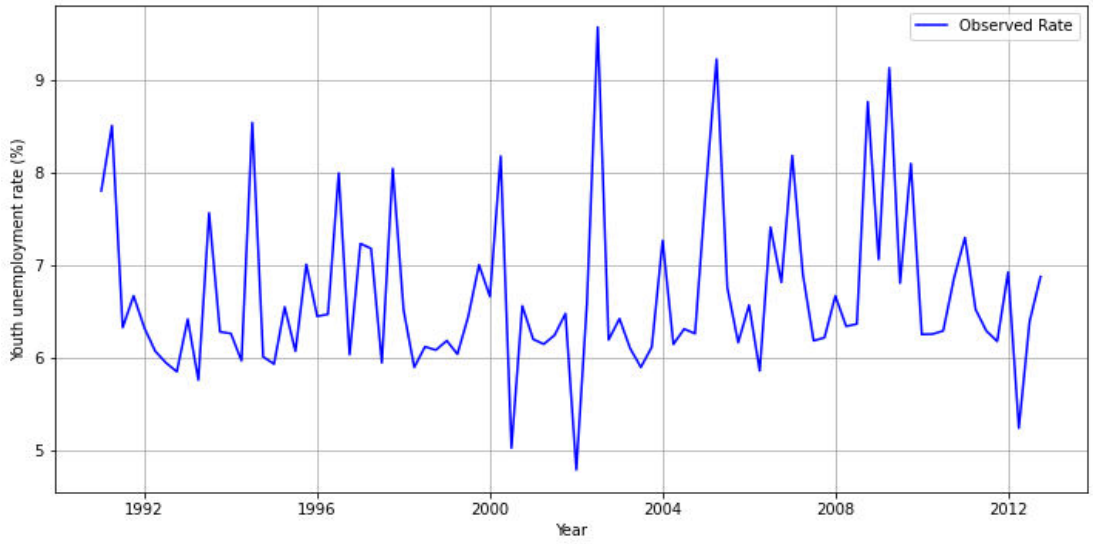

Figure 4.1: Time series plot of observed youth unemployment rates before the devolution period.

From Figure 4.1, it can be seen that the series centers around 7 %, which is closer to the average of the observed rates. The series maintains a fairly steady level over time, that is there is no increasing trend and its variance also seem to remain fairly stable throughout.

c). Autocorrelation Function (ACF)

Figure 4.2 presents the ACF plot up to lag 20, illustrating the estimated autocorrelations between the observed youth unemployment rates and their lagged values. The blue vertical lines represent the estimated autocorrelation coefficients at different lags, with the horizontal thresholds bounded in blue indicating 95% confidence interval (CI) on the estimated autocorrelations. The plot was also used at the model identification stage to identify the order of the moving average component, q and autoregressive components, p in the model.

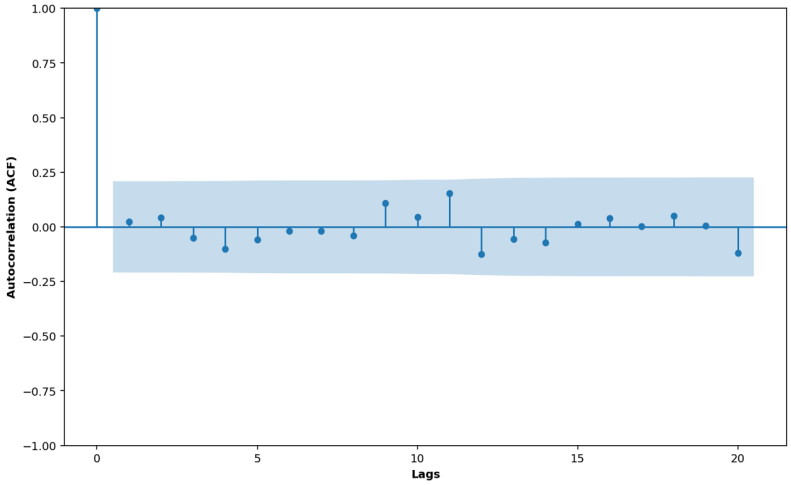

Figure 4.2: ACF plot of the youth unemployment rates before devolution.

From Figure 4.2, all the autocorrelations from lag 1-20 does not exceed the significance bounds at 95% CI. Since there is no significant spike, that is beyond the 95% CI significance bounds, it confirmed a weak relationship between the current rate and its past random errors. This also implies that there is a weak Moving Average (MA) component in the ARIMA model.

d). Partial Autocorrelation Function (PACF)

Figure 4.3 presents the PACF plot up to lag 20, illustrating the estimated partial autocorrelations between the rates and their lagged values.

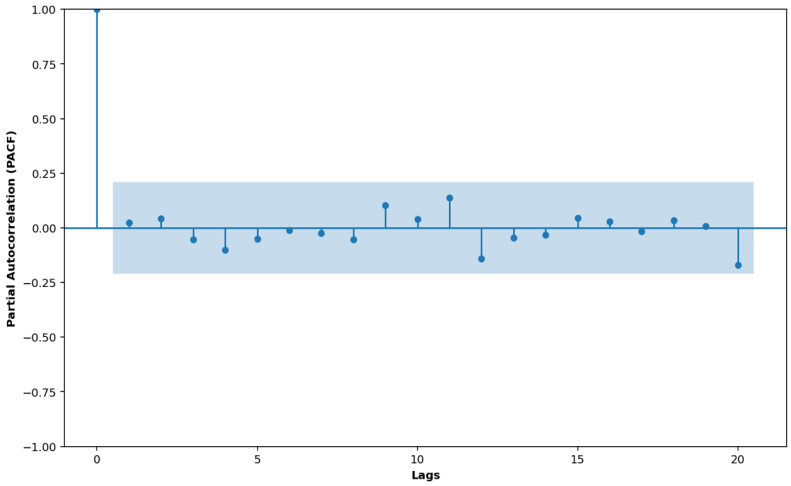

Figure 4.3: PACF plot of the youth unemployment rates before devolution.

From Figure 4.3, all the partial autocorrelations in the plot are within the significance bounds at 95% CI, this implies that there is likely no autoregressive component in the time series, which means that there is a weak relationship between the current youth unemployment rate and its past rate.

e). Seasonality

The aim was to check for strong autocorrelations at seasonal lags, indicating the presence of seasonality component in the series.

Seasonal ACF

The seasonal autocorrelation function (ACF) plot for the stationary data on youth unemployment rates before devolution was examined. The estimated seasonal autocorrelations up to lag 20 are shown in Figure 4.4.

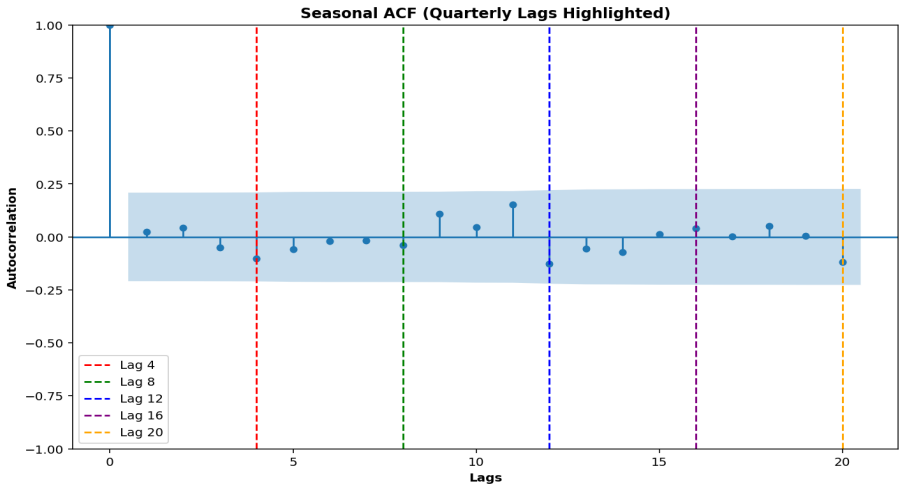

Figure 4.4: Seasonal ACF plot of the observed youth unemployment rates before devolution.

From Figure 4.4, it is clear that most autocorrelations fall within the 95% confidence bounds, indicating weak serial dependence across most lags. However, a notable spike occurs at lag 4, corresponding to a quarterly interval, which suggests the presence of a seasonal pattern repeating every four time periods. The absence of significant autocorrelations at higher seasonal lags (8, 12, 16 and 20) implies that this seasonal effect is limited to the first seasonal cycle. The significant autocorrelation at lag 4 indicates that the series exhibits a quarterly seasonal moving average component that should be accounted for in the model specification.

Seasonal PACF

Figure 4.5 displays the Seasonal PACF of the stationary youth unemployment rate series, with quarterly lags (4, 8, 12, 16 and 20) highlighted. The seasonal PACF measures the correlation between observations at different lags after controlling for the effects of intermediate lags, helping to identify potential autoregressive (AR) terms in a time series model.

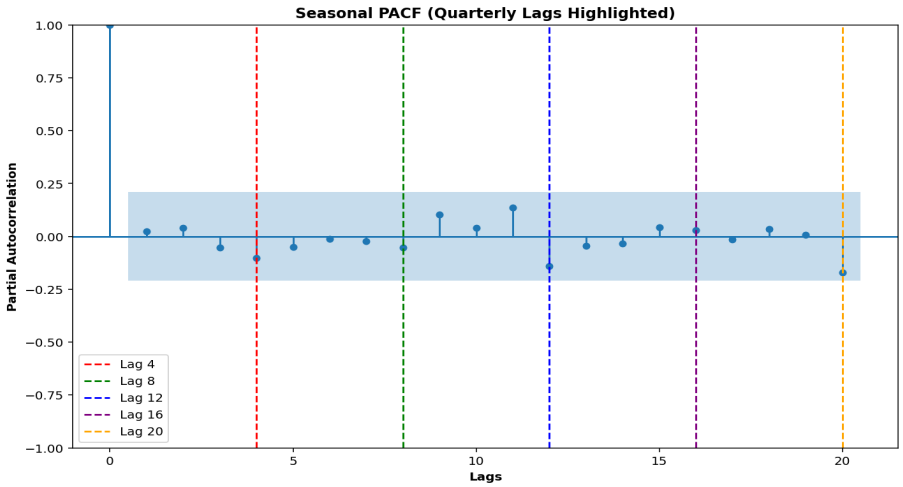

Figure 4.5: Seasonal PACF plot of the observed youth unemployment rates before devolution.

From Figure 4.5, most partial autocorrelations lie within the 95% confidence bounds, indicating weak direct relationships across most lags. However, a notable spike at lag 4 suggests a statistically significant partial autocorrelation. This implies that the series exhibits a seasonal pattern repeating every four periods—likely corresponding to a quarterly effect. The lack of significant spikes at higher seasonal lags (8, 12, 16 and 20) indicates that the seasonal influence is strongest at the first seasonal lag. The PACF plot supports including an autoregressive term at lag 4 to adequately capture the seasonality behavior of the time series.

4.3.2 Model Identification

Here, the plausible values of the moving average and autoregressive components were identified. The main tools used were Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots. Several candidate models were selected based on the initial identification of p, d, q, P, D and Q.

Since all autocorrelations and partial autocorrelations in the ACF plot (Figure 4.2) and the PACF plot (Figure 4.3) beyond lag 0 fall within the 95% confidence interval (within the blue confidence band), it indicates that the series has no significant correlation and behaves like white noise. This means the data is already stationary with no meaningful dependence between observations. This implies that $p = q = d = 0$, which results into an ARIMA(0,0,0). Hence ARIMA(0,0,0) which is fully random with no predictable structure forms the most appropriate representation of the non-seasonal component in the model. However, for com-

pleteness, ARIMA(0,0,1) and ARIMA(1,0,0) can also be tested to confirm the best fit using the AIC value.

We then found the values of P and Q. The delays that are multiples of 4 in Figure 4.4 and Figure 4.5 were noticed, which means the sequence is 4, 8, 12 and 20, since $s = 4$. The seasonal ACF plot in Figure 4.4 show a noticeable spike at lag 4, but is less prominent at higher lags, which are multiples of 4. So we selected $Q = 1$ to be the order of the seasonal Moving Average component. Also, looking at the seasonal PACF plot in Figure 4.5, there is a notable spike at lag 4 and less prominent in other lags. Therefore, the study suggested P=1 to be the order of the seasonal Autoregressive component. There was no seasonal differencing, hence D=0. This suggested that (P, D, Q)=(1,0,0), (0,0,1) and (1,0,1).

4.3.3 Model Comparison and Selection

The Akaike Information Criterion (AIC) value was used at this stage, keeping in mind that given two models, the one with the lowest AIC value is the most preferred. The possible candidate seasonal ARIMA models for youth unemployment rates before devolution are presented in Table 4.2.

Table 4.2: Comparison of the selected seasonal ARIMA models based on their AIC values

Model	**AIC value**	**Std error**
ARIMA(0,0,1)(0,0,1)[4]	236.933	0.117
ARIMA(0,0,0)(1,0,0)[4]	234.887	0.117
ARIMA(0,0,0)(0,0,1)[4]	234.746	0.117
ARIMA(1,0,0)(1,0,0)[4]	237.071	0.117
ARIMA(0,0,0)(1,0,1)[4]	239.977	34.10

Looking at the AIC column in Table 4.2, an ARIMA(0,0,0)(0,0,1)[4] model with a non-zero mean has the lower AIC value of 234.746. This indicates that it performs better than other proposed candidate models, it was therefore selected. ARIMA(0,0,0)(0,0,1)[4] model is a purely seasonal moving average model that captures quarterly patterns in a time series without any trend or non-seasonal components. Since the non-seasonal parameters p, d, and q are all zero, the data is already stationary and not influenced by immediate past values or short-term shocks. In essence, this model represents a situation where quarterly fluctuations are explained primarily by seasonal effects recurring every year, making it suitable for stable quarterly data on youth unemployment rates with consistent seasonal behavior.

4.3.4 Estimated Model Parameters

Here, the parameters of the selected ARIMA(0,0,0)(0,0,1)[4] model with non zero mean were estimated using R programming. The parameter estimates of an ARIMA(0,0,0)(0,0,1)[4] model computed are presented in Table 4.3.

Table 4.3: Parameter estimates of ARIMA(0,0,0)(0,0,1)[4] model.

Parameter	**Estimate**	**Std error**
$\hat{\mu}$	6.6716	0.1195
$\hat{\Theta}$	-0.1254	0.0831

From an ARIMA(0,0,0)(0,0,1)[4] model with a non-zero mean indicated in Table 4.3 implies that the current rate is influenced by the past random errors from the same quarter. Substituting the estimated parameters from Table 4.3 into Equation (3.18), yielded:

$$Y_t = 6.6716 + (-0.1254)e_{t-4} + e_t \tag{4.1}$$

Which translates to;

$$Y_t = 6.6716 - 0.1254e_{t-4} + e_t \tag{4.2}$$

where;

Y_t represents youth unemployment rate at time t

e_{t-4} accounts for the idea that random errors from the same season in previous cycles influence the current youth unemployment rate.

e_t is the usual white noise process.

4.3.5 Model Diagnostic Checking

In this sub section, the study examined how well the selected model fits the data, that is if the residuals of the selected model are white noise, that is if they are uncorrelated, independently and identically distributed (*iid*).

The ACF plot of residuals was examined, looked at the Normal Q-Q plot of residuals, a histogram of residuals overlaid with a normal curve and conducted the Ljung-Box test.

The results are presented in Figures 4.6, 4.7, 4.8 and Table 4.4.

a). ACF Plot of Residuals

Figure 4.6 presents the Autocorrelation Function (ACF) plot, which illustrates the correlation between the model residuals and their lagged values. It was used

to determine whether the residuals (differences between observed and predicted values) exhibit autocorrelation, which would indicate a violation of the independent and identically distributed (*iid*) assumption in the model residuals.

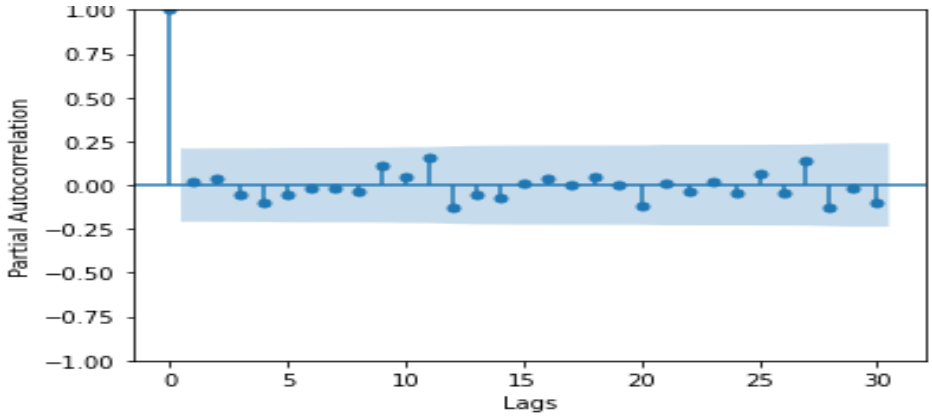

Figure 4.6: ACF plot of residuals.

From Figure 4.6, all the estimated autocorrelations of the residuals for lags 1-30 are within the 95% significance bounds. Since all the lags with autocorrelations does not fall outside these bounds, it suggests that there is no significant relationship between the residuals, which led to a conclusion that they are independent and identically distributed.

b). Normal Q-Q Plot of Residuals

Figure 4.7 presents a Quantile-Quantile plot, which was used to assess the normality assumption of residuals from the fitted model. The plot compared the quantiles of the residuals against the theoretical quantiles of a normal distribution.

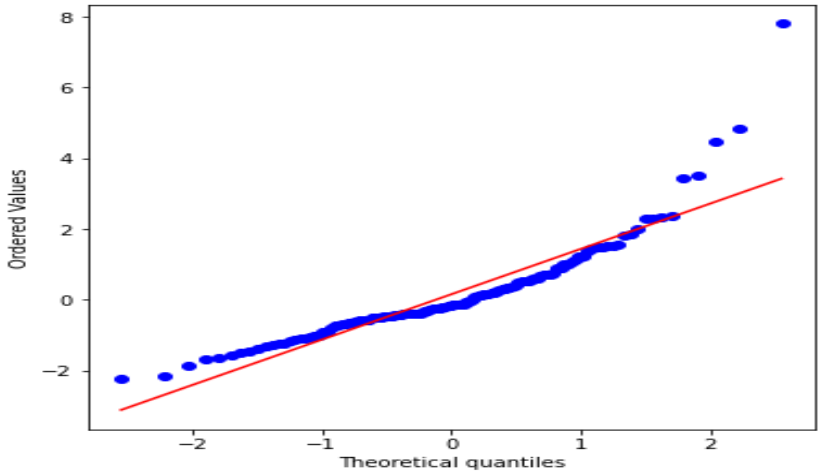

Figure 4.7: Normal Q - Q plot of residuals.

It is clear from Figure 4.7 that the points align closely along the reference line, indicating that the residuals are normally distributed.

c). Histogram of Residuals

Figure 4.8 presents the histogram of residuals obtained from the fitted model. The histogram provided a visual assessment of the distribution of residuals, helping to evaluate whether they follow a normal distribution with mean zero.

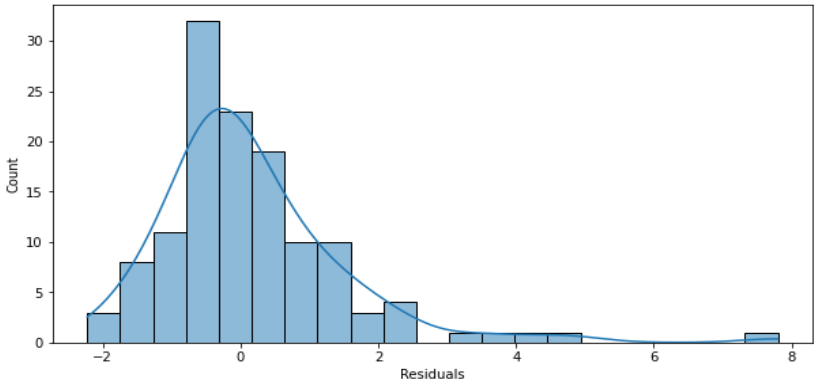

Figure 4.8: Histogram of residuals overlaid with a normal curve.

It appears from Figure 4.8 that the histogram is symmetric and bell-shaped, which indicates that the residuals are normally distributed with an average value estimated at about zero, which is in line with the study assumption that $e_t \sim N(0, \sigma_e^2)$.

d). Ljung-Box Test

This test was used to evaluate whether the model residuals exhibit significant autocorrelation at different lag levels. The following hypotheses were tested at $\alpha = 0.05$.

H_0: The residuals are white noise

Vs

H_a: The residuals are not white noise.

Table 4.4 presents the Ljung-Box test statistics and their corresponding p-values for lag 1 - 10. Stopping at lag 10 is a practical choice to detect autocorrelation while maintaining statistical reliability, avoiding overfitting and ensuring the test remains meaningful for model diagnostics.

Table 4.4: Ljung-Box test results for model residuals.

Lag (k)	Test statistic	P-value
1	0.045972	0.830228
2	0.201101	0.904340
3	0.431557	0.933649
4	1.347016	0.853350
5	1.638560	0.896547
6	1.671012	0.947337
7	1.700811	0.974533
8	1.834562	0.985670
9	2.891880	0.968428
10	3.070372	0.979719

The results in Table 4.4 shows that the p-values across different lags are greater than 0.05, hence we fail to reject the null hypothesis indicating that the residuals are white noise, suggesting that the model sufficiently captures the data before devolution, validating its adequacy.

4.4 ARIMA-Intervention Model

The primary objective was to quantify the effect of devolution so as to find out if it had a statistically significant effect on youth unemployment rates by evaluating the significance of the intervention parameter.

The following hypotheses were tested at 5% significance level.

$$\boldsymbol{H_0}\text{: } \boldsymbol{\omega^{(k)} = 0} \quad \text{Vs} \quad \boldsymbol{H_a}\text{: } \boldsymbol{\omega^{(k)} \neq 0}$$

The computed parameters of ARIMA(0,0,0)(0,0,1)[4] model with devolution intervention are presented in Table 4.5.

Table 4.5: Estimated parameters of ARIMA(0,0,0)(0,0,1)[4] model incorporating devolution intervention.

Variable	Parameter	Estimate	Std error	z-value	P - value	Significance
Intercept	$\hat{\mu}$	6.4982	0.189	34.36	0.000	Significant
Devolution	$\omega^{(k)}$	0.2004	0.010	19.83	0.000	Significant
SMA(1)	$\Theta^{(k)}$	0.2127	0.104	2.038	0.042	Significant
MSE	$\sigma^{(k)}$	1.4543	0.157	9.256	0.000	Significant

According to the fitted ARIMA-Intervention model results presented in Table 4.5, devolution, which was implemented in 2013, had a considerable effect (statistically significant) on the dependent variable of youth unemployment rates. Devolution,

considered as the explanatory variable, is associated with higher youth unemployment rates as indicated by the positive coefficient of the intervention effect parameter (0.2004, $p < 0.001$).

Rewriting Equation (3.39) using the estimated model parameters in Table 4.5, we obtained;

$$Y_t = 6.4982 + 0.2004I_t + 0.2127e_{t-4} + e_t \tag{4.3}$$

where;

Y_t represents youth unemployment rates,

I_t is the occurrence of devolution,

e_{t-4} accounts for the idea that random errors from the same season in previous cycles influence the current youth unemployment rate, and

e_t is the white noise process.

4.4.1 Model Diagnostic Checking for ARIMA - Intervention Model

Model suitability was diagnosed through residual assumption test for adherence to white noise. A histogram overlaid with a normal curve, the Ljung-Box test and a normal Q-Q plot were used.

The results are presented in Figure 4.9, Figure 4.10 and Table 4.6.

a). Histogram of Residuals

Figure 4.9 presents the histogram of residuals obtained from the fitted model. The histogram was used to evaluate whether the residuals follow a normal distribution.

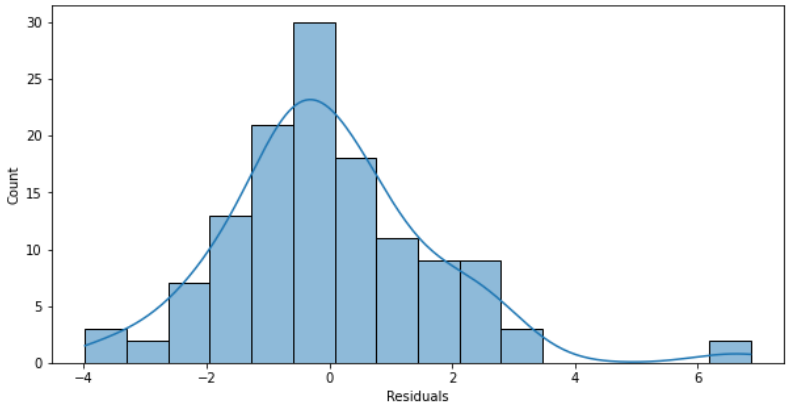

Figure 4.9: Histogram of residuals for ARIMA-Intervention model.

Figure 4.9 suggests that the residuals are likely normally distributed, with a mean approximately zero.

b). Normal Q-Q Plot of Residuals

Figure 4.10 presents a Quantile-Quantile plot, which was used to assess the normality assumption of residuals from the fitted model. The plot was used to compare the quantiles of the residuals against the theoretical quantiles of a normal distribution.

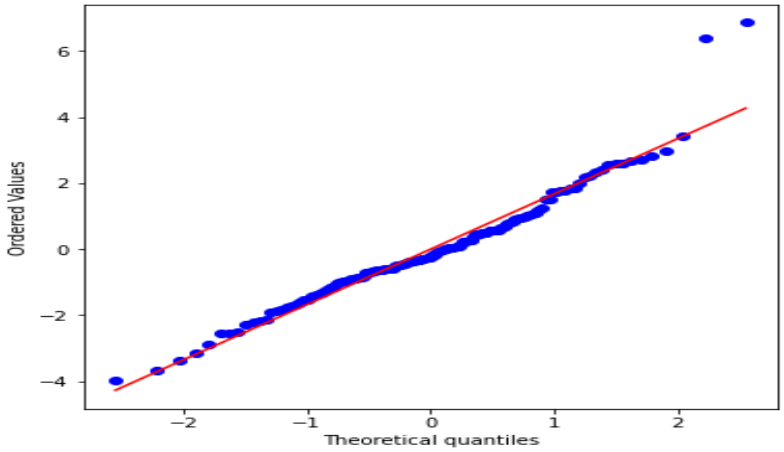

Figure 4.10: Q-Q plot of residuals for ARIMA-Intervention model.

In Figure 4.10, almost all the plotted points in the Q-Q plot perfectly fall along the theoretical normal line, which implies that the residuals follows a normal distribution.

c). Ljung - Box Test

This test was used to evaluate whether the residuals exhibit significant autocorrelation at different lag levels, that is if they are white noise.
Tested hypotheses at 5% significance level:

H_0: The residuals are white noise

Vs

H_a: The residuals are not white noise.

Table 4.6 presents the Ljung - Box test results for lags 1 to 10.

Table 4.6: Ljung - Box test results for model residuals, for white noise assumption.

Lag (k)	Test statistic	P-value
1	0.011993	0.912796
2	0.556732	0.757020
3	1.356709	0.715713
4	1.358038	0.851454
5	1.367066	0.927882
6	1.383118	0.966877
7	1.685359	0.975191
8	1.861703	0.984962
9	1.954703	0.992164
10	2.239935	0.994148

The results in Table 4.6 shows that the p-values from lag 1 to lag 10 are all greater than 0.05. The study therefore failed to reject the null hypothesis and concluded that the residuals are white noise, which implies that they show no autocorrelation and therefore they are independently and identically distributed (iid).

4.5 Forecasted Youth Unemployment Rates (2023–2030)

This section presents the forecast of youth unemployment rates in Kenya for a period of 8 years, using the fitted Autoregressive Integrated Moving Average – Intervention model.

Table 4.7 displays the forecasted quarterly youth unemployment rates from Q_1 2023 to Q_4 2033, along with 95% confidence intervals.

Table 4.7: Quarterly point forecasts for youth unemployment rates with their corresponding 95% confidence intervals.

Date	Rate	Lower 95 CI	Upper 95 CI	Date	Rate	Lower 95 CI	Upper 95 CI
2023 Q_1	13.61	10.34	17.93	2027 Q_1	14.73	9.162	23.69
2023 Q_2	13.68	10.22	18.32	2027 Q_2	14.81	9.118	24.05
2023 Q_3	13.75	10.11	18.70	2027 Q_3	14.88	9.076	24.40
2023 Q_4	13.82	10.01	19.07	2027 Q_4	14.95	9.036	24.75
2024 Q_1	13.89	9.917	19.44	2028 Q_1	15.03	8.997	25.10
2024 Q_2	13.95	9.831	19.81	2028 Q_2	15.10	8.961	25.46
2024 Q_3	14.02	9.751	20.17	2028 Q_3	15.18	8.925	25.81
2024 Q_4	14.09	9.676	20.53	2028 Q_4	15.25	8.891	26.17
2025 Q_1	14.16	9.605	20.88	2029 Q_1	15.33	8.859	26.52
2025 Q_2	14.23	9.539	21.24	2029 Q_2	15.40	8.828	26.88
2025 Q_3	14.30	9.476	21.59	2029 Q_3	15.48	8.798	27.24
2025 Q_4	14.37	9.417	21.94	2029 Q_4	15.56	8.769	27.60
2026 Q_1	14.45	9.361	22.29	2030 Q_1	15.64	8.741	27.97
2026 Q_2	14.52	9.307	22.64	2030 Q_2	15.71	8.715	28.33
2026 Q_3	14.59	9.257	22.99	2030 Q_3	15.79	8.689	28.70
2026 Q_4	14.66	9.208	23.34	2030 Q_4	15.87	8.664	29.06

From Table 4.7, the point forecasts indicated a gradual but consistent rise in youth unemployment rates, beginning at approximately 13.61% in Q_1 2023 and reaching 15.87% by Q_4 2032. A 95% confidence interval accompanied each point forecast, which widens over time from a range of approximately [10.34%, 17.93%] in early 2023 to [8.66%, 29.06%] by the end of 2032, reflecting increasing variability in long-term projections. During the period 2023 to 2026, youth unemployment rates are projected to remain relatively stable. The broader confidence intervals in later years suggests the potential for either a significant decline or moderate increase in the rates depending on future policy implementations and economic developments. This increasing trend underscores the need for early, targeted policy interventions to address structural challenges associated with youth labor market and mitigate the risk of worsening youth unemployment in future.

Figure 4.11 details the observed youth unemployment rates with the eight-year point forecasts for the period 2023 to 2032, as presented in Table 4.7. The shaded region presents the 95% confidence interval on the forecasted rates, indicating the expected range within which the future youth unemployment rates are likely to fall.

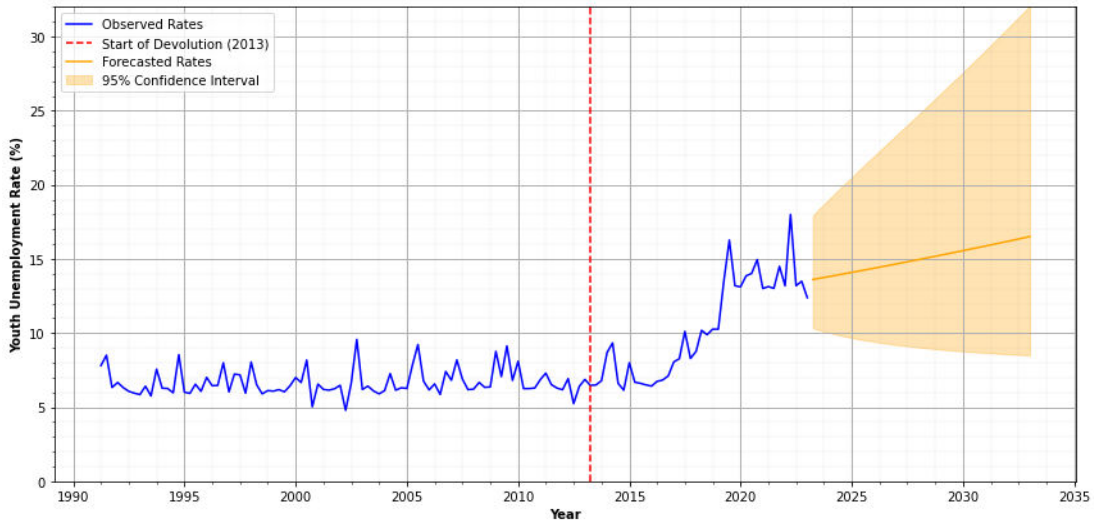

Figure 4.11: The observed historical youth unemployment rates and the 8 year forecasts (2023 - 2030)

From Figure 4.11, the observed youth unemployment rates up to the end of 2022, indicated by the blue line, exhibits seasonal fluctuations and a soft upward trend, suggesting persistent youth unemployment pressure in the labor market even after implementation of devolution governance. The forecasted rates (orange line) extending from 2023 onwards indicates a steady increase in youth unemployment rates, rising from approximately 13.6% in Q_1 2023 to about 15.86% by Q_4 2032. The shaded (orange) region surrounding the forecast line represents the 95% confidence interval bounds, which quantifies the variance associated with the predicted values. This rising trend highlights how the variance forecasts increases as a result of unpredictable macroeconomic or policy changes.

MAPE of the model was approximated to be 10.89%, suggesting that on average, the forecasted rates deviates from the observed rates by 10.89%. This is generally considered a good and acceptable level of forecasting performance for the time series data, particularly in econometric modeling contexts, where noise and variability are relatively high.

4.6 Comparative Analysis

Having evaluated the effect of devolution on youth unemployment rates, comparative analysis was conducted to assess the differences in youth unemployment rate trends before and during the devolution periods. This analysis used both the quantitative forecasting results and qualitative assessment to offer a comprehensive understanding of the rates during the two periods.

4.6.1 Evaluating the Effect of Devolution on Youth Unemployment Rates Using Counterfactual Trends

The pre-intervention model was employed to forecast counterfactual rates for the devolution period. These forecasts were used as a baseline for comparison, enabling a qualitative assessment of the observed deviations and thereby facilitating deeper insights of the effect of devolution.

Figure 4.12 illustrates youth unemployment rates for the period 1991–2022, with a red-dotted vertical red line marking the onset of devolution in 2013. This visual comparison highlighted potential shifts attributable to the implementation of devolution.

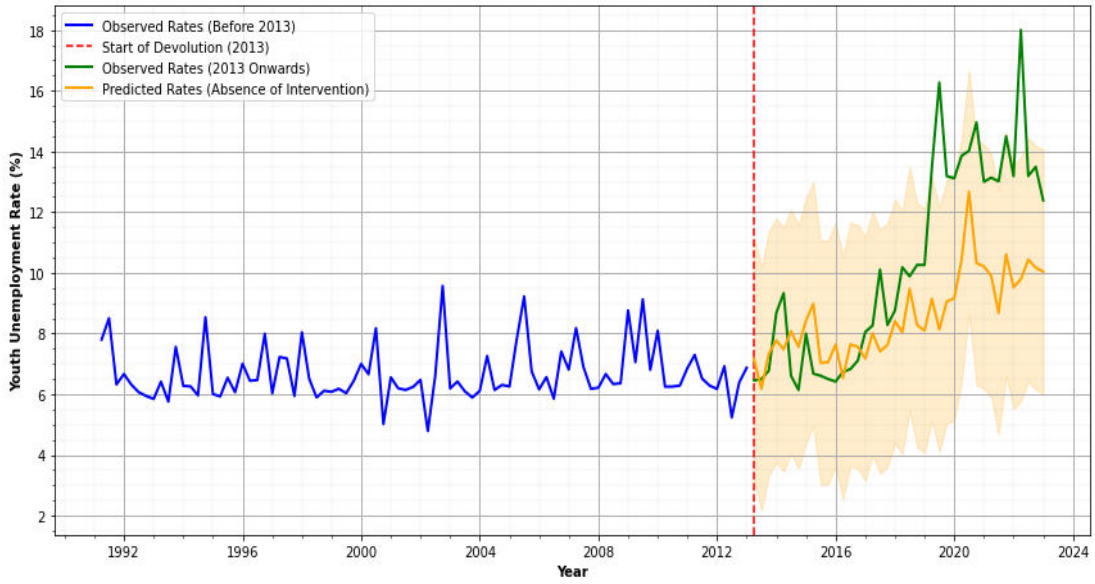

Figure 4.12: Observed and predicted youth unemployment rates with the 2013 intervention point.

From Figure 4.12, the blue line represents the observed youth unemployment rates before devolution, while the green line shows the observed rates during the devolution period. The shaded orange region represents the associated 95% confidence interval, illustrating the range within which the predicted values would likely fall suppose devolution was not implemented.

The intersection points mark the transitions between periods when devolution appeared to be outperforming expectations (green line observed below the orange line) and under-performing expectations (green line observed above the orange line). The initial intersection occurred shortly after 2013. At this point, the observed and predicted rates were almost similar (approximately 6% - 7%), rep-

resenting the moment before the effect of devolution became apparent. At the second intersection (2015 - 2016), the observed youth unemployment rates were temporarily below the predicted rates. This could indicate a brief period where devolution appeared to be working as intended. The third intersection occurred at around 2016 - 2017, after which the observed rates crossed above the predicted line, which marked the beginning of a sustained period where the observed rates significantly exceeded the counterfactual rates.

The fact that the green line (observed rates) spends significantly much of its time above the orange line (predicted rates), suggests that devolution likely had an overall negative effect on youth employment outcomes as compared to what would have been expected suppose it was not implemented.

This unexpected rise contradicts the expectations, as devolution aimed to create more job opportunities. For instance, macroeconomic factors such as exchange rate volatility, inflation and slow economic growth might have strained policy implementation, making it harder for youths to find employment despite the efforts of devolution. Structural factors such as skills mismatch between youths and labour market demands, might have also played a role, limiting the effectiveness of devolution in addressing the issue of youth unemployment. These findings implies that while devolution has had a notable effect, additional measures are needed to optimize its benefits in scaling youth employability.

4.6.2 Youth Unemployment Rates before and during Intervention

To further assess the effect of devolution on youth unemployment rates, youth unemployment rates before and after the devolution periods were compared. The analysis incorporated a counterfactual scenario to estimate what the youth unemployment trend was expected to be suppose devolution was not implemented.

Figure 4.13 presents a flow chart illustrating the trend in youth unemployment rates before and during the devolution periods, alongside the counterfactual estimate.

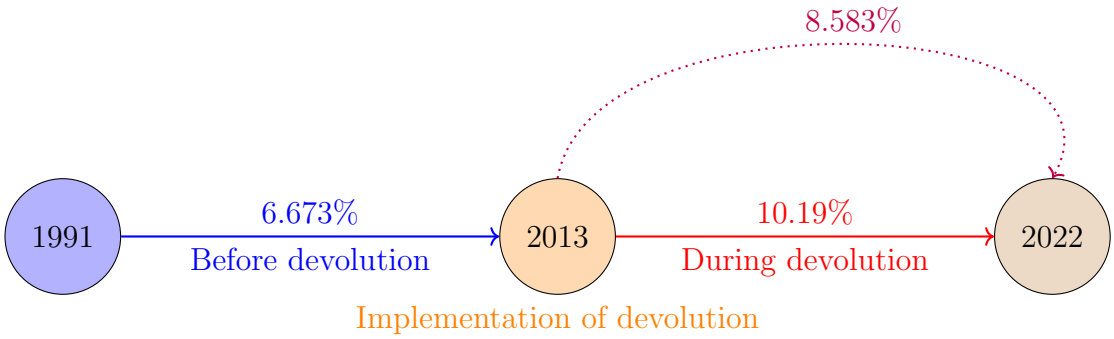

Figure 4.13: A flow chart showing youth unemployment rates before and during the devolution period.

As illustrated in Figure 4.13, the average youth unemployment rate before devolution was approximately 6.67%. Following the introduction of devolution, this rate increased to about 10.19%. In comparison, the counterfactual scenario assuming devolution was not implemented suggests an average youth unemployment rate of approximately 8.583%, which is notably lower than the observed rate during the devolution period. The estimated absolute effect of devolution on youth unemployment is therefore approximately 1.607%. In relative terms, this reflects an increase of 24.09%, which implies that the increase in youth unemployment during the devolution period was statistically significant and likely attributable to devolution (which aligns with the tested significance of the impact parameter, ω in the fitted ARIMA - Intervention model, *see Table 4.5*).

This study findings are consistent with other research as reviewed in the literature, that emphasizes the inefficiencies of devolved governance. Monari *et al.* (2020) and Sam *et al.* (2020) demonstrated that Kenya is experiencing a *"jobless growth,"* where macroeconomic progress does not transfer into sufficient youth employment opportunities. Nzau (2014) argued that mismanagement of devolved funds hampered economic growth. These similarities imply that governance issues and structural inefficiencies within devolved units may be the cause of the observed increase in youth unemployment during the devolution period.

However, these findings contradicts those of Nimrod (2016), who stated that devolution has improved socioeconomic performance in the Nakuru East constituency. This significant difference may be as a result of methodological differences, that is, in the current study, time series modeling represents the nationwide outcomes on youth unemployment, which may be more affected by systemic problems such as

weak institutional capacity, political instability and corruption, whereas localized studies focuses on context specific improvements.

Methodologically, the current study builds on earlier studies as reviewed in the literature (Adesegun *et al.* 2020; Adenomon, 2017; Karlson Javed, 2016) on unemployment forecasting, by incorporating devolution as a structural policy intervention in the ARIMA model. This demonstrates how powerful time series intervention modeling can be used in reflecting how significant governance reforms affects labor markets.

4.7 Fitted Dynamic Adaptive Intervention Model

4.7.1 Testing for Normality Assumption

Histograms shown in Figure 4.14 were utilized in assessing the distributional properties of the residuals. It displays the distribution of residuals for each macroeconomic variable included in the analysis. This visual assessment was useful in determining whether the residuals approximated a normal distribution, an important requirement for the validity of parametric inference procedures.

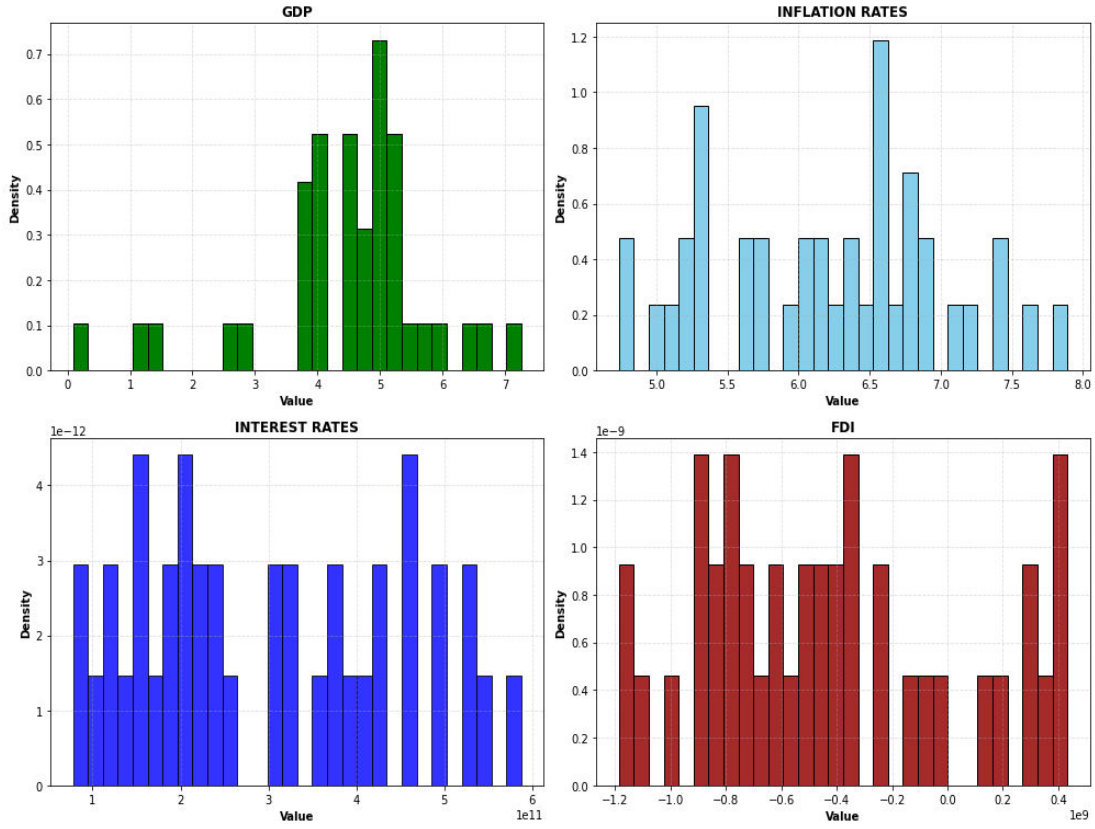

Figure 4.14: Histogram of residuals for each economic variable

Figure 4.14 revealed that the residuals for all the economic variables considered

in this study do not follow a normal distribution. The figures show a noticeably irregular shape with bars that are widely spaced, indicating skewness and potential outliers. These deviations from normality suggests that the underlying parametric model assumptions may not be fully met for these variables. This non-normality could affect the reliability of inference based on standard parametric tests, and therefore, the alternative non-parametric approach (Quantile regression) was considered for more accurate interpretations.

4.7.2 Estimated Dynamic Effect of Devolution on Youth Unemployment Rates

This subsection presents the estimated values of the dynamic effect and associated uncertainty values within the adaptive intervention model over time. Table 4.8 summarizes the time-varying estimates of the devolution's effect, $\hat{\beta}_t$ alongside the corresponding uncertainty measures, P_t, prior to the inclusion of any control for economic variables. These results provides insights into the evolving strength and reliability of the effect of devolution across quarterly periods from 2013 to 2022.

Table 4.8: Estimated dynamic effect of devolution on youth unemployment, $\hat{\beta}_t$ and uncertainty, P_t values before accounting for macroeconomic variables

Year	$\hat{\boldsymbol{\beta}}_{\boldsymbol{t}}$	$\boldsymbol{P_t}$	**Year**	$\hat{\boldsymbol{\beta}}_{\boldsymbol{t}}$	$\boldsymbol{P_t}$
$2013Q_1$	0.0000	1.0000	$2018Q_1$	31.4089	0.0270
$2013Q_2$	6.3309	0.0910	$2018Q_2$	32.6384	0.0270
$2013Q_3$	9.2463	0.0502	$2018Q_3$	33.9881	0.0270
$2013Q_4$	12.8936	0.0376	$2018Q_4$	34.8887	0.0270
$2014Q_1$	15.9732	0.0322	$2019Q_1$	36.8037	0.0270
$2014Q_2$	17.8070	0.0297	$2019Q_2$	39.3023	0.0270
$2014Q_3$	19.7072	0.0284	$2019Q_3$	40.2618	0.0270
$2014Q_4$	21.7226	0.0278	$2019Q_4$	41.1485	0.0270
$2015Q_1$	22.8083	0.0274	$2020Q_1$	41.7020	0.0270
$2015Q_2$	23.9648	0.0272	$2020Q_2$	42.1126	0.0270
$2015Q_3$	24.6978	0.0271	$2020Q_3$	42.5047	0.0270
$2015Q_4$	25.3189	0.0271	$2020Q_4$	43.3101	0.0270
$2016Q_1$	26.1034	0.0270	$2021Q_1$	44.4532	0.0270
$2016Q_2$	26.3605	0.0270	$2021Q_2$	45.9379	0.0270
$2016Q_3$	26.6356	0.0270	$2021Q_3$	48.6893	0.0270
$2016Q_4$	27.2962	0.0270	$2021Q_4$	50.5596	0.0270
$2017Q_1$	27.7180	0.0270	$2022Q_1$	52.3062	0.0270
$2017Q_2$	28.7609	0.0270	$2022Q_2$	53.0962	0.0270
$2017Q_3$	29.0482	0.0270	$2022Q_3$	53.9012	0.0270
$2017Q_4$	29.6492	0.0270	$2022Q_4$	54.5411	0.0270

From Table 4.8, the parameter $\hat{\beta}_t$, which represents the estimated time varying effect of devolution on youth unemployment rates at time t. The positive increasing values of $\hat{\beta}_t$ suggests that devolution had a growing effect on youth unemployment, i.e youth unemployment rates increased during devolution. The uncertainty values, P_t decreased gradually, suggesting that the model gains confidence in its estimates over time.

4.7.3 Estimated Effect of Macroeconomic Variables

A quantile regression model was fitted using the estimated dynamic effect $\hat{\beta}_t$ in Table 4.8 as the dependent variable and the macroeconomic variables (GDP, Inflation Rates(I), interest income on the Kenya's public debt (IR) and Foreign Direct Investment (FDI)) as the explanatory variables.

The specified quantile regression model based on Equation (3.41) was expressed as:

$$\beta_t = \gamma_0(\tau) + \gamma_1(\tau)X_{t1} + \gamma_2(\tau)X_{t2} + \gamma_3(\tau)X_{t3} + \gamma_4(\tau)X_{t4} + \eta_t \tag{4.4}$$

where;

τ is the selected quantile level ($\forall\ 0 \leq \tau \leq 1$)

β_t is the dynamic effect of devolution on youth unemployment computed using the Kalman filter equations, $\gamma_0(\tau)$ is the intercept value calculated at τ quantile level, $\gamma_1(\tau)$ is the effect of GDP growth at τ, X_{t1} represents the GDP growth at time t, $\gamma_2(\tau)$ is the effect of inflation rate at τ, X_{t2} represents Inflation rate at time t, $\gamma_3(\tau)$ is the effect of interest income at τ, X_{t3} represents interest income on the Kenya's public debt at time t, $\gamma_4(\tau)$ is the effect of Foreign Direct Investment (FDI) at τ, X_{t4} represents Foreign Direct Investment (FDI) at time t, and η_t represents the residuals (and η_t is distribution free).

Parameter estimation was done by the use of optimization method (linear programming) so as to minimize the quantile loss function associated with the deviations. The Pseudo R-squared values were also estimated at different quantile levels so as to evaluate the fitness of the fitted models. The p-vales of the model parameters were estimated so as to make inferences on the effect of the control macroeconomic variables on the estimated state dynamic effect of devolution.

The following hypotheses were tested at 5% significance level:

H_0: Estimated parameter is statistically insignificant

V/s

H_a: Estimated parameter is statistically significant

Table 4.9 presents the results of the quantile regression estimating the effect of key macroeconomic variables on the dynamic effect of devolution at the lower quartile (25th percentile). The model included GDP, Inflation Rates, interest income on the Kenya's public debt and Foreign Direct Investment (FDI) as predictors.

Table 4.9: Quantile regression of the estimated dynamic effect of devolution at 25th percentile level ($\tau = 0.25$)

Variable	**Parameter**	**Estimate**	**Std err**	***t* statistic**	**P - value**	**Significance**
Intercept	γ_0	0.8810	6.686	0.132	0.896	Insignificant
GDP	γ_1	0.4798	0.491	0.976	0.336	Insignificant
Inflation	γ_2	0.2473	0.862	0.287	0.776	Insignificant
Interest	γ_3	0.8807	0.072	12.271	0.000	Significant
FDI	γ_4	0.0041	0.198	0.021	0.983	Insignificant

Pseudo $R - squared$: 0.7791

From Table 4.9, the following interpretations were made based on the p-values at 5% significance level and direction of coefficients.

Intercept: The intercept is statistically insignificant ($p = 0.896 > 0.05$), with a positive value. The positive intercept suggests a baseline increase in the time-varying effect at the lower tail of the distribution.

Gross Domestic Product (GDP): The coefficient for GDP indicates a positive effect and is statistically insignificant ($p = 0.336 > 0.05$), suggesting that an increases in GDP was associated with an increase in the dynamic effect of devolution at the lower quartile level.

Inflation rate: The coefficient is statistically insignificant ($p = 0.776 > 0.05$) and positively affects the lower-end devolution effect.

Interest income: Interest income was statistically significant ($p \approx 0.000 < 0.05$) and had a positive effect on the dynamic effect of devolution. This implies that a one-unit increase in interest income increased the estimated dynamic effect of devolution by approximately 0.8807 units at 25th percentile level.

Foreign Direct Investment (FDI): The FDI coefficient is positive and statistically insignificant ($p = 0.983 > 0.05$), indicating that increased FDI may have contributed positively to the dynamic effect of devolution at the 25th percentile level.

Table 4.10 presents the results of the quantile regression estimating the effect of GDP, Inflation Rates, interest income and Foreign Direct Investment (FDI) as predictors on the dynamic effect of devolution at the median level (50th percentile).

Table 4.10: Quantile regression of the estimated dynamic effect of devolution at 50th percentile level ($\tau = 0.50$)

Variable	**Parameter**	**Estimate**	**Std err**	***t* statistic**	**P - value**	**Significance**
Intercept	γ_0	5.2551	3.543	1.483	0.147	Insignificant
GDP	γ_1	0.2251	0.301	0.748	0.460	Insignificant
Inflation	γ_2	0.5249	0.517	1.015	0.317	Insignificant
Interest	γ_3	0.7944	0.040	20.050	0.000	Significant
FDI	γ_4	0.0470	0.127	0.369	0.714	Insignificant

Pseudo $R - squared$: 0.8314.

From Table 4.10, the following interpretations were made based on the p-values at 5% significance level and direction of coefficients.

Intercept: The intercept is statistically insignificant ($p = 0.147 > 0.05$), with a positive value, which reflects the baseline level of the dynamic effect of devolution when all predictors are zero.

Gross Domestic Product (GDP): The coefficient for GDP is positive and statistically insignificant ($p = 0.460 > 0.05$).

Inflation rate: The coefficient for inflation is positive and statistically insignificant ($p = 0.284 > 0.05$).

Interest income: Interest income was statistically significant ($p \approx 0.000 < 0.05$) and had a positive effect on the dynamic effect of devolution. This implies that a one-unit increase in interest income increases the estimated dynamic effect of devolution by approximately 0.7944 units at 50th percentile level.

Foreign Direct Investment (FDI): The FDI coefficient is positive and statistically insignificant ($p = 0.714 > 0.05$), indicating that an increase in FDI may have contributed to an increase in the dynamic effect of devolution at the 50th percentile level.

Table 4.11 presents the results of the quantile regression estimating the effect of key macroeconomic variables on the time varying impact of devolution at the upper quartile (75th percentile).

Table 4.11: Quantile regression of the estimated dynamic effect of devolution at 75th percentile level ($\tau = 0.75$)

Variable	Parameter	Estimate	Std err	*t* statistic	P - value	Significance
Intercept	γ_0	6.9301	3.487	1.988	0.055	Insignificant
GDP	γ_1	0.2183	0.289	0.755	0.455	Insignificant
Inflation	γ_2	0.4147	0.526	0.788	0.436	Insignificant
Interest	γ_3	0.7899	0.036	21.897	0.000	Significant
FDI	γ_4	0.1049	0.121	0.866	0.392	Insignificant

Pseudo $R - squared$: 0.8676

From Table 4.11, the following interpretations were made based on the p-values at 5% significance level and direction of coefficients.

Intercept: The intercept value is statistically insignificant ($p = 0.055 > 0.05$) in the fitted model.

Gross Domestic Product (GDP): The coefficient for GDP is positive and statistically insignificant ($p = 0.455 > 0.05$).

Inflation rates: The coefficient for inflation is positive and not statistically significant ($p = 0.436 > 0.05$) in the model.

Interest income: Interest income was statistically significant ($p \approx 0.000 < 0.05$) and had a positive effect on the dynamic effect of devolution. This suggests that higher interest income was associated with higher dynamic effect of devolution on youth unemployment rates at 75th percentile level.

Foreign Direct Investment (FDI): The FDI coefficient is positive and statistically insignificant ($p = 0.392 > 0.05$) in the model.

4.7.4 Estimated Dynamic Effect of Devolution Based on the Fitted Models

Table 4.12 summarizes the time-varying estimates of the devolution's effect on youth unemployment rates, $\hat{\beta}_t$ was computed using the fitted adaptive intervention models across the 25th, 50th and 75th percentiles from 2013 to 2022.

Table 4.12: Estimated dynamic effect of devolution on youth unemployment rates, $\hat{\beta}_t$ at 25th, 50th and 75th percentile levels

Year	**25%**	**50%**	**75%**	**Year**	**25%**	**50%**	**75%**
$2013Q_1$	10.94	14.87	15.33	$2018Q_1$	31.96	33.66	33.97
$2013Q_2$	11.80	15.69	16.12	$2018Q_2$	32.30	33.68	34.12
$2013Q_3$	13.91	17.63	18.06	$2018Q_3$	33.99	35.06	35.59
$2013Q_4$	16.10	19.65	20.07	$2018Q_4$	33.70	34.89	35.45
$2014Q_1$	15.97	19.56	19.96	$2019Q_1$	35.50	36.62	37.20
$2014Q_2$	16.90	20.43	20.85	$2019Q_2$	37.11	38.16	38.78
$2014Q_3$	19.21	22.54	23.02	$2019Q_3$	37.07	38.22	38.85
$2014Q_4$	19.21	22.55	23.09	$2019Q_4$	39.08	40.40	41.15
$2015Q_1$	18.83	22.22	22.81	$2020Q_1$	38.90	40.58	41.43
$2015Q_2$	20.53	23.78	24.40	$2020Q_2$	39.85	41.79	42.75
$2015Q_3$	20.57	23.83	24.46	$2020Q_3$	40.26	42.50	43.57
$2015Q_4$	21.50	24.68	25.32	$2020Q_4$	44.01	45.80	46.90
$2016Q_1$	23.05	26.10	26.75	$2021Q_1$	44.45	45.86	46.91
$2016Q_2$	22.02	25.18	25.82	$2021Q_2$	45.62	46.58	47.58
$2016Q_3$	22.63	25.78	26.26	$2021Q_3$	49.35	49.63	50.61
$2016Q_4$	24.10	27.16	27.48	$2021Q_4$	49.58	49.66	50.56
$2017Q_1$	24.15	27.25	27.41	$2022Q_1$	52.31	52.31	53.08
$2017Q_2$	25.76	28.76	28.76	$2022Q_2$	51.87	52.08	52.72
$2017Q_3$	26.23	29.02	29.05	$2022Q_3$	53.90	54.09	54.60
$2017Q_4$	27.61	29.99	30.15	$2022Q_4$	56.87	56.95	57.34

The results in Table 4.12 revealed a steadily increasing trend in the time varying effect of devolution on youth unemployment rates across all the considered distributional quantiles (25th, 50th and 75th), indicating that the negative effect of devolution on youth unemployment intensified consistently from 2013 to 2022. At the onset of devolution in 2013, the estimated effects were relatively moderate, but by 2022, they had more than quintupled, with values converging around 56 to 57 across all percentiles. This underscores the reality that devolution, rather than providing immediate employment relief, has progressively contributed to higher unemployment rates among youth regardless of their prior efforts to curb the problem, even after accounting for the key macroeconomic factors e.g. GDP, inflation rates, interest income on the Kenya's public debt and Foreign Direct Investment. This could be due to challenges in implementing effective labor market policies, inappropriate resource allocation or economic uncertainties associated with the transition to devolved governance. These results highlights need for redesigning the policies at the county level so as to include youth employment strategies.

4.7.5 Model Selection

To obtain the most appropriate quantile regression model in assessing the time-varying effect of devolution on youth unemployment rates in Kenya, a comparison of model performance across different quantile levels was conducted based on the estimated Pseudo R-Squared values.

Table 4.13 presents Pseudo R-Squared values obtained from quantile regression models estimated at the 25th, 50th and 75th quantile levels. These values served as measures of model fitness, indicating how well each quantile model explained the variation in the dynamic effect at different points of its conditional distribution.

Table 4.13: Comparison of Pseudo R-Squared for quantile regression at 25th, 50th and 75th percentile levels

Model	Quantile level	Pseudo R-squared
1.	25th percentile	0.7791
2.	50th percentile	0.8314
3.	75th percentile	0.8676

As shown in Table 4.13, the Pseudo R-squared values increased with the quantile level, suggesting that the model's explanatory power improved at higher points in the conditional distribution of the dynamic effect of devolution. The model at the 25th percentile explained approximately 77.91% of the variation, while the 50th and 75th percentile models achieved 83.14% and 86.76% respectively. This upward trend indicates that the effect of the macroeconomic control factors becomes more pronounced toward the upper quantiles. Consequently, the model at the 75th percentile level may offer the best fit overall, and it was therefore selected.

Substituting the estimated parameters from Table 4.11 into the specified model in Equation (4.4) at $\tau = 0.75$;

$$\hat{\beta}_t = 6.9301 + 0.2183GDP_{t1} + 0.4147I_{t2} + 0.7899IR_{t3} + 0.1049FDI_{t4} + \eta_t \quad (4.5)$$

where; $\hat{\beta}_t$ is the estimated conditional dynamic effect of devolution on youth unemployment at the 75th percentile, GDP_{t1} is the Gross Domestic Product at time t, I_{t2} is inflation rate at time t, IR_{t3} represents interest income at time t, FDI_{t4} represents Foreign Direct Investment (FDI) at time t and η_t represents the deviations of actual values from the estimated conditional values.

CHAPTER FIVE

CONCLUSION AND RECOMMENDATIONS

5.1 Introduction

This chapter summarizes key study findings in Section 5.2 and provides recommendations based on the results from both the ARIMA-Intervention and Dynamic Adaptive Intervention models in Section 5.3.

5.2 Conclusion

This study set out to model and evaluate the effect of devolution on youth unemployment rates in Kenya using an ARIMA–Intervention modeling framework, supported by Dynamic Adaptive Intervention Modeling (DAIM). The fitted ARIMA(0,0,0)(0,0,1)[4] model incorporating a ramp intervention function demonstrated that devolution had a statistically significant effect on youth unemployment rates at 5% level of significance. This confirms that ARIMA–Intervention modeling is appropriate for assessing structural policy effects on labor market time series.

Forecast projections for the period 2023–2030 showed a continued increasing trend in youth unemployment if current conditions persist, which suggests that without strategic employment-centered reforms under devolution, youth unemployment levels are likely to remain high. The forecasts highlight the need for proactive interventions to reverse the projected increasing trend.

Comparative analysis findings disclosed that the average youth unemployment rate stood at 6.67% during the pre-devolution period, which increased to 10.19% during the devolution period. The counterfactual rate was estimated to be 8.583%, representing the average rate suppose devolution was not implemented, which indicates that the observed increase was unlikely by chances, which confirms that devolution had a significant effect on youth unemployment rates.

The DAIM results revealed that the negative effect of devolution on youth unemployment progressively increased over time, indicating a gradual worsening trend. The quantile regression modeling further revealed that interest income on Kenya's public debt resulted into a negative influence across all the considered distributional quantiles, which underscores the role of fiscal policies in shaping the labor market outcomes. This implies that the strength of devolution in job creation has also been weakened by macroeconomic pressures and structural financial restrictions.

5.3 Recommendations

The following recommendations aim to address the key challenges based on the findings of this study. These suggestions are designed to be practical and aligned with the best practices that drive purposeful improvements in modeling the effect of devolution on youth unemployment rates in Kenya.

The study assumed a linear association between youth unemployment rates and devolution, which may not have fully captured the composite socio-economic effects of devolution on youth unemployment. Therefore, the study recommends considering alternative machine learning methods, such as Random Forests, which can capture non-linear relationships and interactions between variables. Additionally, approaches like Elastic Net Regression could provide flexibility in modeling the data without the need for strict linearity assumptions.

The study recommends that future studies should consider expanding the age limit of youth to capture individuals aged 25 – 35 in addition to 15 - 24 bracket, which might ensure a more comprehensive understanding of youth unemployment in Kenya. This broader definition would capture the employment challenges faced by older youths, who often experience distinct difficulties such as delayed education to employment transitions, additional financial responsibilities and age-related discrimination.

REFERENCES

[1] Adesegun, O. I., Mathew, O., & Omotola, S. B. (2020). Modeling unemployment rates in Nigeria using time series approach. *Asian Journal of Mathematical Sciences, 4*(2). Federal University of Agriculture.

[2] Adenomon, N. O. (2017). Modeling and forecasting unemployment rates in Nigeria using ARIMA model. *Statistics Unit, Department of Mathematical Sciences, Nasarawa State University,* 2(1), 525–531.

[3] Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis: Forecasting and control* (1st ed.). Holden-Day.

[4] Box, G. E. P., & Tiao, G. C. (1975). Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical Association, 70*(349), 70–79. https://doi.org/10.1080/01621459.1975.10480264

[5] Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (1976). *Time series analysis: Forecasting and control.* Holden-Day.

[6] Brockwell, P. J., & Davis, R. A. (2002). *Introduction to time series and forecasting* (2nd ed.). Springer.

[7] Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research, 33*(2), 261–304.

[8] Collins, L. M., Murphy, S. A., & Strecher, V. (2004). The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): New methods for more potent health interventions. *American Journal of Preventive Medicine, 32*(5), 112–118.

[9] Didiharyono, D., & Syukri, M. (2020). Forecasting with ARIMA model in anticipating open unemployment rates in South Sulawesi. *International Journal of Scientific & Technology Research, 9*(3).

[10] Faguet, J. P. (2014). Decentralization and governance. *World Development, 53,* 2–13. https://doi.org/10.1016/j.worlddev.2013.01.002

[11] Hall, S. (2017). *Youth employment in Kenya: Literature review.* British Council.

[12] Koenker, R., & Hallock, K. F. (2001). Quantile regression: An introduction. *Journal of Economic Perspectives, 15*(4), 43–56.

[13] Hivos. (2016). *The status of the creative economy in East Africa (Ubunifu Report).*

[14] Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (2nd ed.). OTexts. Available at: *https://otexts.com/fpp2/*

[15] International Labour Organization. (2020). *Global employment trends for youth 2020.* International Labour Office.

[16] International Labour Organization. (2022). *Statistics on youth: Concepts and definitions.* International Labour Office.

[17] International Labour Organization. (2024). *Global employment trends for youth 2024.* International Labour Office.

[18] ILOSTAT. (2023). *Youth labour statistics database.* International Labour Organization.

[19] Karlson, S., & Javed, F. (2016). Modeling and forecasting unemployment rate in Sweden using various econometric measures (Master's thesis). *Örebro University, School of Business.*

[20] Katumo, M. D., & Maingi, J. (2020). Relationship between youth unemployment and economic growth in Kenya. *School of Economics, Kenyatta University, 8*(5), 255–267.

[21] Kenya Institute for Public Policy Research & Analysis (KIPPRA). (2012). *Youth unemployment in Kenya: Its nature and covariates.* KIPPRA. https://repository.kippra.or.ke/bitstreams/7751a4b6-0c07-4b5e-a6ced67c2e5e217d/download

[22] Kenya Institute for Public Policy Research and Analysis (KIPPRA). (2023). *Empirical review of youth employment policies and their impact in Kenya* (Policy Brief).

[23] Kilonzo, S. (2011). *Governance and business in a devolved system in Kenya.*

[24] Kenya National Bureau of Statistics (KNBS). (2019). *Kenya population and housing census: Volume III, distribution of population by age and sex.*

[25] Kenya National Bureau of Statistics (KNBS). (2019). *Quarterly labour force report: Quarter 1 (January–March).*

[26] Kenya National Bureau of Statistics (KNBS). (2022). *Quarterly labour force report: Quarter 4 (October–December).*

[27] Lam, D. (2014). *Youth bulges and youth unemployment* (IZA World of Labor No. 26). Institute of Labor Economics (IZA).

[28] Lidia, C. (2011). Decentralization in Africa: Scope, motivations and impact on service delivery and poverty. *Working Paper 020.*

[29] Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika, 65*(2), 297–303.

[30] Monari, F. N., Tinega, R., & Agasa, L. O. (2020). Modeling economic determinants of youth unemployment using multiple linear regression: A case study of Kenya. *International Journal of Statistics and Applied Mathematics, 5*(6), 7–11.

[31] Munyao, C. (2019). *Youth unemployment and its implications on development in Africa: The case of Kenya* (Master's thesis). University of Nairobi.

[32] Mwakalila, E. (2023). Income inequality: A recipe for youth unemployment in Africa. *SN Business & Economics*, 3(1), 15. https://doi.org/10.1007/s43546-022-00394-0

[33] Navaei, N. (2023). ARIMA vs GARCH (time series forecasting). *LinkedIn.* Available at: *https://www.linkedin.com/pulse/arima-vs-garch-time-series-forecasting-mahdi-navaei*

[34] Ngigi, S., & Busolo, D. N. (2019). Devolution in Kenya: The good, the bad and the ugly. *Public Policy and Administration Research, 9*(6).

[35] Nimrod, K. K. (2016). The role of devolution on economic development in Kenya: A case of Nakuru East Sub-county. *Research Paper.* Maasai Mara University, School of Business and Economics.

[36] Nzau, M. (2014). The effects of devolved funds on economic growth in Kenya: An empirical investigation (1993–2012) (Master's thesis). *University of Nairobi.*

[37] Oates, W. E. (1999). An essay on fiscal federalism. *Journal of Economic Literature, 37*(3), 1120–1149. https://doi.org/10.1257/jel.37.3.1120

[38] Okuom, J., Obange, N., & Odhiambo, S. (2023). Effect of per capita income on youth unemployment in Kenya. *International Journal of Economics, 8*(2), 1–18. https://doi.org/10.47604/ijecon.2118

[39] Otoi, S. S., & Pokhariyal, G. P. (2016). *Modelling economic determinants of youth unemployment in Kenya.* Journal of Emerging Trends in Economics and Management Sciences, 7(1), 31-38.

[40] Putnam, R. J. (1993). *Making democracy work: Civic traditions in modern Italy.* Princeton University Press.

[41] Ray, M., Ramasubramanian, V., Kumar, A., & Rai, A. (2014). Application of time series intervention modeling and forecasting cotton yield. *Statistics and Applications, 12*(1–2), 61–70.

[42] Riechi, B. (2019). *Youth unemployment and security in Kenya: The case of Kwale County (2010–2018)* (Master's thesis). United States International University–Africa.

[43] Rondinelli, D. A., Nellis, J. R., & Cheema, G. S. (1983). *Decentralization in developing countries: A review of recent experience* (World Bank Staff Working Paper No. 581). World Bank.

[44] Sam, S. O., Manene, M. M., Kipchirchir, I. C., & Pokhariyal, G. P. (2020). Cointegration analysis of youth unemployment in Kenya. *International Journal of Statistics and Applied Mathematics, 5*(3), 129–133.

[45] Sharahiley, S. (2020). Forecasting unemployment rate in Saudi labour market in 2030: The story of downward resistance. *Journal of Organizational Behaviour Research, 5*(1), 330–349.

[46] Smoke, P. (2003). Decentralisation in Africa: Goals, dimensions, myths and challenges. *Public Administration and Development, 23*(1), 7–16. https://doi.org/10.1002/pad.255

[47] Song, S. M., & Wittes, J. (2018). Modeling dynamic treatment regimes with time-varying effects in longitudinal studies. *Biometrics, 74*(3), 803–812.

[48] United Nations, Department of Economic and Social Affairs. (2013). *Definition of youth.* United Nations.

[49] United Nations Development Programme (UNDP). (2022). *Devolution: From inception to conception and beyond.* UNDP Kenya.

[50] UNICEF. (2023). *Why devolution must work for children, youth, women, and marginalized groups in Kenya.* UNICEF Kenya.

[51] Wamalwa, F. M. (2009). *Youth unemployment in Kenya: Its nature and covariates* (KIPPRA Discussion Paper No. 103). Kenya Institute for Public Policy Research and Analysis.

[52] World Bank. (2023). *World development indicators: Youth unemployment, ages 15–24.* The World Bank.

[53] Yu, D. (2013). Youth unemployment in South Africa since 2000 revisited. *Economic Working Documents,* 4–13.

APPENDICES

Appendix I: Interpolated Quarterly Youth Unemployment Rates (%)

Year	Q_1	Q_2	Q_3	Q_4
1991	7.803	8.506	6.324	6.667
1992	6.319	6.069	5.943	5.847
1993	6.415	5.757	7.564	6.278
1994	6.260	5.965	8.539	6.008
1995	5.930	6.547	6.068	7.007
1996	6.446	6.468	7.994	6.032
1997	7.231	7.178	5.943	8.043
1998	6.512	5.895	6.117	6.081
1999	6.182	6.037	6.440	7.002
2000	6.659	8.177	5.022	6.558
2001	6.197	6.145	6.242	6.475
2002	4.787	6.596	9.572	6.191
2003	6.421	6.099	5.895	6.114
2004	7.265	6.144	6.309	6.258
2005	7.826	9.225	6.756	6.162
2006	6.566	5.857	7.409	6.813
2007	8.183	6.901	6.182	6.215
2008	6.667	6.337	6.363	8.764
2009	7.060	9.130	6.804	8.095
2010	6.250	6.254	6.288	6.861
2011	7.297	6.520	6.288	6.175
2012	6.924	5.238	6.392	6.873
2013	6.461	6.497	6.780	8.686
2014	9.334	6.609	6.144	7.999
2015	6.683	6.609	6.502	6.419
2016	6.739	6.830	7.111	8.055
2017	8.266	10.110	8.283	8.761
2018	10.184	9.888	10.264	10.261
2019	13.474	16.273	13.189	13.117
2020	13.858	14.021	14.958	13.002
2021	13.136	13.014	14.500	13.188
2022	18.004	13.193	13.494	12.389

Appendix II: Interpolated Quarterly Macroeconomic Indicators

Year	GDP Growth (%)	Inflation Rate (%)	FDI (KES)	Interest Income (KES)
2013	3.845	5.742	-9.08E+08	7.76E+10
2013	4.067	5.978	-8.83E+08	8.55E+10
2013	4.423	6.286	-8.24E+08	1.07E+11
2013	4.788	6.597	-7.62E+08	1.29E+11
2014	4.904	6.777	-7.65E+08	1.26E+11
2014	4.990	6.821	-6.95E+08	1.36E+11
2014	5.149	6.844	-5.65E+08	1.61E+11
2014	5.060	6.735	-5.00E+08	1.62E+11
2015	4.931	6.604	-4.46E+08	1.59E+11
2015	4.961	6.587	-3.62E+08	1.78E+11
2015	4.692	6.471	-3.82E+08	1.80E+11
2015	4.518	6.405	-3.76E+08	1.92E+11
2016	4.410	6.375	-3.52E+08	2.10E+11
2016	4.033	6.201	-4.00E+08	2.01E+11
2016	3.964	6.605	-5.99E+08	2.07E+11
2016	3.987	7.058	-7.74E+08	2.22E+11
2017	3.858	7.429	-9.90E+08	2.23E+11
2017	3.897	7.890	-1.16E+09	2.40E+11
2017	4.031	7.450	-1.19E+09	2.46E+11
2017	4.401	6.660	-1.09E+09	2.61E+11
2018	5.090	6.039	-9.13E+08	3.09E+11
2018	5.348	5.189	-8.48E+08	3.13E+11
2018	5.613	4.735	-7.43E+08	3.32E+11
2018	5.349	4.788	-7.02E+08	3.30E+11
2019	5.309	4.959	-6.02E+08	3.51E+11
2019	5.248	5.118	-5.08E+08	3.69E+11
2019	5.006	5.182	-4.61E+08	3.69E+11
2019	3.907	5.297	-2.34E+08	3.98E+11
2020	2.571	5.286	-6.93E+07	4.03E+11
2020	1.357	5.340	1.28E+08	4.20E+11
2020	0.085	5.362	3.09E+08	4.31E+11
2020	1.112	5.606	4.33E+08	4.68E+11
2021	2.750	5.675	3.85E+08	4.64E+11
2021	4.466	5.785	3.59E+08	4.67E+11
2021	6.458	6.041	4.05E+08	4.98E+11
2021	7.264	6.166	3.04E+08	4.96E+11
2022	6.766	6.596	1.94E+08	5.28E+11
2022	5.928	6.845	-6.54E+06	5.27E+11
2022	5.354	7.235	-1.37E+08	5.53E+11
2022	4.865	7.670	-2.46E+08	5.88E+11

Appendix III: Modeling and Empirical Analysis Codes

R Procedure for Fitting ARIMA - Intervention Model

```
# Required packages library(readxl)
library(forecast)
library(tseries)
library(TSstudio)
library(lmtest)
# Import data
file_path <- "C ://Users/SAMMY/Desktop/mydata12.xlsx"
data <- read_excel(file_path, sheet = 1)
ts_data <- ts(data$Rate, start = c(1991, 1), frequency = 4)
#Define intervention date
intervention <- ifelse(time(ts_data) >= 2013, time(ts_data) - 2013, 0)
# Model fitting
fit <- Arima(ts_data, order = c(0, 0, 0), seasonal = c(0, 0, 1), xreg = interven-
tion)
# Print the model
summary(fit)
coeftest(fit)
```

R Procedure for obtaining Counterfactual Rate

```
# Required packages library(CausalImpact)
library(zoo)
library(readxl)
# Load data
data <- read_excel("C ://Users/SAMMY/Desktop/mydata12.xlsx")
# Ensure no missing values
data <- na.omit(data)
# Create a proper Date column
data$Date <- as.Date(paste(data$Year, (data$Quarter - 1) * 3 + 1, "01", sep =
"-"))
# Sort data to ensure chronological order
data <- data[order(data$Date), ]
# Convert to a time-series object using "zoo"
data_zoo <- zoo(data$Rate, order.by = data$Date)
# Define pre and post periods
pre_period <- c(min(data$Date), as.Date("2012-12-31"))
post_period <- c(as.Date("2013-01-01"), max(data$Date))
# Run Causal Impact analysis
impact < - CausalImpact(data_zoo, pre_period, post_period)
# Print summary
summary(impact)
plot(impact)
```