MODELLING THE EFFECT OF COVID – 19 MORTALITY SHOCK ON FERTILITY RATE IN KENYA USING NEGATIVE BINOMIAL REGRESSION MODEL

Linus Omina Wakhu

A Thesis Submitted to the School of Natural Sciences in Partial Fulfillment of the Requirements of the Award of the 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 no other than the indicated sources and support and has not been presented elsewhere for a degree or any other award

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

Wakhu Linus Omina

SES/G/01-70033/2020

CERTIFICATION

The undersigned certify that they have read and hereby recommend for acceptance of Masinde Muliro University of Science and Technology a research thesis entitled, "Modelling the Effect of COVID-19 Mortality Shock on Fertility Rate in Kenya using Negative Binomial Regression Analysis"

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

Dr. Everlyne Akoth Odero

Department of Mathematics

Masinde Muliro University of Science and Technology

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

Dr. Drinold Mbete

Department of Mathematics

Masinde Muliro University of Science and Technology

COPYRIGHT

This thesis is 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 or anyone in full or in parts except for short extracts in fair dealings, for research or private study, critical scholarly reviewing or disclosure with acknowledgement, with written permission of the Director Directorate of Postgraduate Studies on behalf of both the author and Masinde Muliro University of Science and Technology.

DEDICATION

I dedicate this work to my family, whose unwavering support and encouragement have been the foundation of my journey.

ACKNOWLEDGEMENTS

I am thankful to the Almighty God for the gift of life, good health, care, protection, knowledge and wisdom that He granted me on the journey to pursue this MSc degree. My deepest gratitude to my supervisors, Dr. Everlyne Akoth Odero and Dr. Drinold Mbete for their continuous support throughout the research work. It is through their intellectual input, resourcefulness, dedication, effort and prudent scientific and moral guidance that I am acquainted with a lot of knowledge and skills in various aspects of my research. I thank the entire staff of Mathematics department for their encouragement and support offered during the entire period of my research. I cannot forget to thank Masinde Muliro University of Science and Technology for permitting me to undertake my MSc in Statistics. My gratitude goes to my dear family, parents and siblings who encouraged me throughout the entire research work.

ABSTRACT

Fertility is known as one of the three primary components that determine the structure, size, and composition of the population of any country, the other two being mortality and migration. Mortality has been shown to have some effects on fertility by various researchers. The world has experienced mortality shocks as a result of various deadly disasters that have previously occurred, from the influenza epidemics of 1918 to COVID-19 that was first reported in 2019. Various studies have been conducted by various researchers on the effects of mortality shocks on fertility but little documentation on effect of mortality shock as a result of COVID-19 on fertility is available. Some of the mortality shocks experienced were as a result of epidemics, natural calamities and violences i.e influenza epidemic, spanish flu, famines, genocides, tsunamis, earthquakes, ebola and severe acute respiratory syndrome virus. More recently the world experienced COVID-19 epidemic which resulted into mortality shock. Few empirical studies have examined the impact of the COVID-19 mortality shock on fertility. This study addresses this gap by modeling the effect of COVID-19–related mortality on fertility rates in Kenya using the Negative Binomial Regression model. The study used secondary data from KDHS. Parameters were etimated using maximum likelihood estimation (MLE), R- software was used for analysis. Results of analysis based on both IRR and p-values revealed that the effect of COVID-19 mortality shock was significant on fertility rate. There was a slightly higher expected count value of 7.488 after COVID-19 compared to 7.146 Before COVID-19. The two values were statistically significant at $\alpha = 0.05$. The dispersion parameters of 1.537 and 1.572 had a minimal difference, indicating that both models captured overdispersion similarly. The significance of this study is that it offers key insights into how mortality shocks like COVID-19 affect fertility trends in Kenya, providing evidence to guide demographic projections, policy decisions, and further research on the links between health crises, reproductive behavior, and population dynamics. The study concluded that the COVID-19 mortality shock had a positive impact on fertility.

TABLE OF CONTENTS

TITLE PAGE i

DECLARATION ii

COPYRIGHT iii

DEDICATION iv

ACKNOWLEDGEMENT v

ABSTRACT vi

TABLE OF CONTENTS vii

LIST OF TABLES x

LIST OF FIGURES xi

LIST OF ABBREVIATIONS AND ACRONYMS xii

BASIC CONCEPTS AND TERMS xiii

CHAPTER ONE: INTRODUCTION **1**

1.1 Background Information 1

1.2 Statement of the Problem 4

1.3 Objectives of Study 5

1.3.1 Main Objective 5

1.3.2 Specific Objectives 6

1.4 Significance of the Study 6

1.5 Justification of the Study 7

1.6 Research Methods 8

1.6.1 Source of Data 8

1.6.2 Negative Binomial Regression 8

1.6.3 Maximum Likelihood Estimation (MLE) 9

1.6.4 Mean Squared Error (MSE) 10

CHAPTER TWO: LITERATURE REVIEW 12

2.1 Introduction 12

2.2 Theoretical perspectives on the mortality-fertility relationship . . . 12

2.3 Empirical evidence from past mortality shocks 13

2.4 COVID-19 and fertility 13

2.5 Methodological approaches and rationale for the Negative Binomial Model 14

2.6 Developed Studies using Poisson and Negative Binomial Regression Model 16

CHAPTER THREE: MODEL DESCRIPTION 22

3.1 Introduction 22

3.2 Model Assumptions 22

3.3 The Negative Binomial Regression Model 22

3.4 Mean and Variance of the Model 23

3.5 Regression Specification 24

3.6 Log-Likelihood Function 24

3.7 Parameter Estimation 24

CHAPTER FOUR: RESULTS AND DISCUSSION 29

4.1 Introduction 29

4.2 Fitting the Negative Binomial Regression model to COVID-19 mortality shock free data set in Kenya and Parameter Estimation . . . 29

4.2.1 The results of the Model before COVID-19 31

4.3 Fitting the Negative Binomial Regression model to data set affected by COVID-19 mortality shock in Kenya and Parameter Estimation 33

4.3.1 The results of the Model after COVID-19 35
4.4 Comparison of the result of the model on KDHS data before and after COVID-19 37
4.4.1 Density Curve of the Negative Binomial Regression model on KDHS datasets before and after COVID-19 37
4.4.2 Result from model fitting on KDHS data before and after COVID-19 compared 39
4.4.3 Estimated Count Before COVID-19 40
4.4.4 Estimated Count After COVID-19 41
4.5 Testing the Hypotheses of μ_a and μ_b 41

CHAPTER FIVE: CONCLUSION AND RECOMMENDATIONS 44

5.1 Introduction 44
5.2 Summary of Findings Based on Study Objectives 44
5.3 Conclusion 45
5.4 Policy Recommendations 45
5.5 Areas for Further Research 46

REFERENCES 46

APPENDICES 53

LIST OF TABLES

Table 1.1: Mean squared error check for Poisson and Negative Binomial models 11

Table 2.1: The dispersion test results for Poisson model 15
Table 2.2: AIC test results for Poisson and Negative Binomial models . 16

Table 4.1: The Table of Extracted explanatory variables before COVID-19 30
Table 4.2: Table of the estimated parameters of the regression model before COVID-19 31
Table 4.3: The Table of Extracted explanatory variables after COVID-19 34
Table 4.4: Table of the Estimated parameters of the Regression model After COVID-19 35

LIST OF FIGURES

Figure 3.1: Visual display of fertility determinants of interest 26

Figure 4.1: Density Curve showing the behavior of fertilty before and after COVID-19 37

Figure 4.2: Visual comparison of estimated fertility counts before and after COVID-19 42

LIST OF ABBREVIATIONS AND ACRONYMS

AIDS	Acquired Immunodeficiency Syndrome
ASFR	Age Specific Fertility Rate
COV	Coronavirus
COVID-19	Coronavirus Disease - 2019
GLM	Generalized Linear Model
HIV	Human Immunodeficiency Virus
IRLS	Iteratively Reweighted Least Squares
IRR	Incidence Rate Ratio
KDHS	Kenya Demographic and Health Survey
KNBS	Kenya National Bureau of Statistics
MLE	Maximum Likelihood Estimation
MSE	Mean Squared Error
NBR	Negative Binomial Regression
NDHS	Nigeria Demographic and Health Survey
NSFG	National Survey of Family Growth
PMF	Probability Mass Function
PRM	Poisson Regression Model
SARS	Severe Acute Respiratory Syndrome
SHHS	Sudan Household Health Survey
TFR	Total Fertility Rate
WHO	World Health Organization

BASIC CONCEPTS AND TERMS

Baby Boom	A period characterized by a significant increase in the number of births.
COVID-19	A communicable respiratory disease caused by the SARS-CoV-2 virus.
Fertility	The actual number of live children born to women within a population.
Fertility Rate	The number of live births per 1,000 women of reproductive age, categorized by specific maternal characteristics.
Generalized Linear Model (GLM)	A flexible generalization of the ordinary least squares regression that relates the distribution of the response variable (expected value) to the systematic portion of the experiment (linear predictors).
Mortality Shock	A sudden, unexpected rise in deaths resulting from a calamity, pandemic, or similar catastrophic event.

CHAPTER ONE

INTRODUCTION

1.1 Background Information

Fertility refers to the actual number of live births and is one of the three fundamental components of population dynamics-alongside mortality and migration-that determine the size, structure, and composition of a country's population [43]. Population growth rates are often viewed as indicators of a nation's economic development [9]. In this context, the recent decline in fertility across many African countries is perceived as an opportunity that may contribute to sustained economic growth and improved living standards [8, 36].

It has been observed that fertility level of a population does not only determine the current size but also has a positive or negative effect on the future economic growth of a population. High fertility and rapid population growth are a hindrance to economic growth and development implying persistent poverty cycle [38].

Davis and Blake [14] suggested two types of factors that mainly affect fertility i.e, the direct and indirect factors. The direct (proximate) determinants have direct connection to fertility, they include; contraceptive use, induced abortion, sterility, woman sexuality and post-partum insusceptibility. The indirect (background) factors influencing fertility include respondents' age, education level, employment status, area of residence, and religion, among others. Mortality shocks have been observed to influence some of the direct determinants of fertility-such as marriage patterns, contraceptive use, and birth spacing-thereby exerting an overall effect on fertility rates. [3, 35, 20].

Various studies [32, 3, 12] have explored fertility in the aftermath of high-mortality disasters. A large number of deadly disasters have previously occurred in the world history. From influenza epidemics to COVID-19, all have resulted into mortality shock.

The Great Finnish Famine which occurred between the year 1866 and 1868 [34] killed many people in Finland, the number was estimated to be more than two hundred thousand people. The birth rates later remarkably increased shortly after the famine ended [42].

The 1994 Rwandan genocide resulted in a significant mortality shock that profoundly affected demographic patterns in the country. Empirical evidence indicates that the genocide had diverse effects on fertility, varying according to the type of violence experienced by women, their age cohort, parity, and the post-conflict time horizon. Notably, there is strong evidence of a child replacement effect, where women who experienced the loss of a child during the genocide demonstrated a higher likelihood of both desiring and having additional children in the five years following the conflict [40]. The a bove findings shows how extreme mortality events can reshape reproductive behavior and fertility results in post-crisis populations.

The Spanish flu that was witnessed in the year 1918 was a very destructive pandemic in modern history and killed 50 million people worldwide. A unique feature of this virus was the high death rate among young adults aged 20–40 years [23]. In 1920, a baby boom occurred in European countries including Norway, Sweden, and the UK. This surge was identified as reflecting the tendency of many couples rushing to wed and then conceive children after surviving the epidemic [28].

The 2004 Indian Ocean tsunami resulted into high mortality shock which killed large shares of the residents of some Indonesian communities [6]. There was a

sustained fertility increase at the aggregate level following the tsunami. Mothers who lost children in the disaster were found to be more likely to bear additional children after the tsunami [16].

A surge in fertility rate was also observed after the Great Kanto Earthquake in Japan that occurred in the year 1923. Experts suggested that victims of the disaster sought motherhood because of child loss in the earthquake [14].

The Severe Acute Respiratory Syndrome (SARS) epidemic which was experienced in the year 2003, resulted into mortality shock that led into a decline in fertility rates. In a study carried out in Taiwan that compared to pre-SARS period, the market share for childbirth health services dropped in medical centers and regional hospitals with reduced cesarean rates during the peak SARS period [25]. Afterwards, fertility rate rose significantly.

The Ebola epidemic in Africa killed about half of infected individuals, spread by means of bodily fluids, and had a very high fatality rate. During the Ebola epidemic, the birth rate declined, but after the announcement of several countries as being Ebola-free, the birth rate temporarily rose [45]. For instance, in Liberia, a sharp decline in birth rates was observed during the first 6 months from the beginning of the Ebola outbreak, whereas a rise was reported for 5 months in the 17 months following the outbreak [30].

COVID-19 is an infectious disease caused by Severe Acute Respiratory Syndrome Coronavirus 2, which was first reported in December 2019 in Wuhan, People's Republic of China. The spread of this disease was extremely rapid throughout the world [22]. The World Health Organization [46] declared COVID-19 as a Public Health Emergency on January 30, 2020. By March 11, 2020, COVID-19 had killed over one million people and hence was declared as a global pandemic by WHO.

The COVID-19 virus spreads primarily through close person-to-person contact, most often via small respiratory droplets produced when an infected person sneezes, coughs, or talks. Transmission can also occur indirectly through contact with contaminated surfaces, followed by touching the face, nose, or eyes with unwashed hands. To prevent the spread, governments worldwide instituted tough public health regulations, including lockdowns, which affected more than seventy five percent of the global population for periods ranging from weeks to months. These measures significantly affected social and reproductive behaviors, as increased number of families spent more time together at home, leading to reports of increased conception rates and a subsequent rise in births in the months following lockdowns [33].

The mentioned cases shows how mortality shocks and related societal disruptions can influence fertility patterns. However, there is minimal documentation on the incorporation of mortality shocks resulting specifically from the COVID-19 pandemic and their effect on fertility rates. This study therefore sought to fill the gap by modeling the effect of COVID-19 mortality shock on fertility rates in Kenya using Negative Binomial Regression model.

1.2 Statement of the Problem

COVID-19 pandemic had far reaching social, economic, and demographic effects across the globe, extending far beyond its immediate health consequences. In Kenya, while high attention was given to the direct impacts of the pandemic, such as mortality, morbidity, and economic disruption, its effect on reproductive behavior and fertility patterns remains insufficiently studied. Fertility choices are strongly influenced by variations in mortality risk, economic uncertainty, and social instability, all of which were highly intensified during the period of the pandemic.

The pandemic therefore resulted into mortality shock, this had the potential to change fertility dynamics through various mechanisms including replacement fertility, delayed childbearing, and changes in family planning choices.

available research on the demographic impacts of COVID-19 has been dominated by studies from high income nations, with relatively low empirical evidence from sub-Saharan Africa. Most of these studies depends on aggregate or cross-sectional data that sometimes ignores local-level diffences and contextual causes special to African populations. Not forgetting, earlier analyses of fertility responses to mortality shocks and pandemics have mostly used models such as Poisson regression, logistic regression, or time-series trend analyses. Even though they are useful, these models sometimes fail to account for over-dispersion, a common feature in fertility data where the variance exceeds the mean, leading to biased or inefficient estimates.

This methodological gap necessitates the need for more robust analytical approaches that is able to handle count data with higher variability. The Negative Binomial Regression model offers a suitable alternative, as it accommodates over dispersion and gives more reliable estimates of fertility outcomes. Therefore, this study sought to fill the empirical and methodological gaps by modeling the effect of COVID-19 mortality shock on fertility rates in Kenya using Negative Binomial Regression model. The study's findings will play a pivotal role as it will enhance understanding of how pandemics influence fertility behavior in low and middle income populations and offer evidence to inform post-pandemic population and health policy planning.

1.3 Objectives of Study

1.3.1 Main Objective

To model the effect of COVID-19 mortality shock on fertility rate in Kenya using Negative Binomial Regression model.

1.3.2 Specific Objectives

(i) To fit Negative Binomial Regression model to COVID-19 mortality shock free data set.

(ii) To fit Negative Binomial Regression model to data set affected by COVID-19 mortality shock.

(iii) To determine the effect of COVID-19 mortality shock on fertility rate in Kenya.

1.4 Significance of the Study

The findings from this study will offer crucial insights into the connection between mortality shocks and fertility trends in Kenya. Understanding how fertility rates behaves due to mortality shock such as the COVID-19 pandemic will assist the government and policymakers expect demographic shifts and design appropriate social, economic, and health interventions.

Demographers will also benefit from this study as it gives empirical evidence that can be employed to improve population projection models, particularly in post shock recovery periods. Statisticians will not be left out as they will find the study valuable for its application of the Negative Binomial Regression model in analyzing fertility data, hence enhancing methodological approaches for handling over-dispersed count data.

Academicians will also use the findings to support further research on the demographic impacts of pandemics and related shocks, while medical practitioners will gain a better understanding of how health crises influence reproductive behavior and population health outcomes. Overall, the study contributes to evidence based planning and a broader understanding of the effects of mortality shocks on fertility

patterns in Kenya.

1.5 Justification of the Study

Comprehending how mortality shocks affect fertility is vital for demographic planning and policy formulation. Despite worldwide attention to the pandemic, its demographic effect, especially on fertility is still under explored in sub-Saharan Africa, including Kenya. Most of the available studies have emphasized health and economic outcomes, overlooking how mortality shocks changes reproductive behavior. Earlier research mostly have used Poisson, OLS, or logistic regression models, which are limited when handling fertility data that exhibit over dispersion. Incorporating such models can result into biased estimates and unreliable inferences. To overcome these limitations, this study applies the Negative Binomial Regression (NBR) model, which introduces a dispersion parameter to account for unobserved heterogeneity, making it more suitable for fertility analysis. By applying the NBR model to Kenya Demographic and Health Survey (KDHS) data, the study offers a more accurate assessment of how COVID-19 mortality shocks influenced fertility rates. The findings will enhance demographic modeling and support evidence based population and health policy in Kenya's post pandemic context.

1.6 Research Methods

1.6.1 Source of Data

This study used secondary data from Kenya Demographic Health Survey (KDHS). Two data sets were used to model the effect of COVID – 19 mortality shock on the fertility rates of Kenya using Negative binomial regression analysis. Before COVID-19 the Kenya Demographic Health Survey of 2018 and after COVID-19 the Kenya Demographic Health Survey of 2022 were used.

1.6.2 Negative Binomial Regression

Negative Binomial Regression model was used to model the effect of COVID - 19 mortality shock on fertility rate in Kenya.

Negative Binomial Regression is a generalized linear model (GLM) form of regression analysis used to model count data that exhibits overdispersion[19, 10, 48]. It is the most appropriate distribution for modeling rates, that are based on the count of events occurring in a certain amount of time where the assumption of Poisson is violated [10, 13]. Negative Binomial Regression and Poisson Regression are similar, but the only difference that exists between them is that the dependent (y) variable is an observed count which follows the Negative Binomial distribution. Accordingly, y possible values are non negative integers: 0,1,2,3, and so on [15]. It provides a better fit by introducing an additional dispersion parameter.

Negative binomial regression is a generalization of the Poisson Regression model that accounts for this overdispersion.

The Negative Binomial distribution is a discrete probability distribution of the number of successes in a sequence of independent and identically distributed Bernoulli trials before a specified (non-random) number of failures (which is denoted as $\frac{1}{k}$) occurs. That is the probability distribution of the number of successes before the

k^{th} failure occurs is a Negative Binomial distribution.

The Negative Binomial Regression model is written as;

$$Y_i \sim NB(\mu_i, \alpha) \tag{1.1}$$

Where;

Y_i is the count dependent variable.

α is the dispersion parameter

μ_i is the expected count for observation i modelled as;

$$log(\mu_i) = X_i\beta \tag{1.2}$$

$X_i\beta$ represents the linear combination of predictors.

β represents the predictors

The Negative binomial regression is similar to Poisson distribution but with a dispersion parameter α.

1.6.3 Maximum Likelihood Estimation (MLE)

Maximum likelihood estimation was used to estimate the parameters (B_s) of the Negative Binomial Regression model.

This was achieved by maximizing a likelihood function so that, under the assumed statistical model, the observed data is most probable. The point in the parameter space that maximizes the likelihood function is called the maximum likelihood estimate [37].

Let $x_1, x_2.....x_n$ be a random sample from a distribution that depends on one or more unknown parameters $\theta_1, \theta_2.....\theta_m$ with probability density function;

$$f(x_i : \theta_1, \theta_2.....\theta_m)$$

Suppose that $(\theta_1, \theta_2.....\theta_m)$is restricted to a given parameter space Ω then;

(i) When regarded as a function of $\theta_1, \theta_2.....\theta_m$ the joint probability density function of $x_1, x_2.....x_n$

$$L(\theta_1, \theta_2.....\theta_m) = \prod_{i=1}^{n} f(x_i : \theta_1, \theta_2.....\theta_m) \qquad (1.3)$$

$L(\theta_1, \theta_2.....\theta_m)$ in Ω is called the Likelihood function

(ii) If $[\mu_1(x_1, x_2.....x_n), \mu_2(x_1, x_2.....x_n), ...\mu_m(x_1, x_2.....x_n)]$ is the $m-tuple$ that maximises the likelihood function, then;

$$\hat{\theta}_i = \mu_i(x_1, x_2.....x_n) \qquad (1.4)$$

is the maximum Likelihood estimator of θ_i for $i = 1, 2, ..., m$.

(iii) The corresponding observed values of the statistic

$$[\mu_1(x_1, x_2.....x_n), \mu_2(x_1, x_2.....x_n), ...\mu_m(x_1, x_2.....x_n)] \qquad (1.5)$$

are called the Maximum Likelihood estimate of θ_i for $i = 1, 2, ..., m$.

1.6.4 Mean Squared Error (MSE)

Mean Squared Error was used to check predictive accuracy of the model i.e how close the predicted fertility counts are to the actual observed counts.

Mean Squared Error is represented as:

$$MSE = \frac{\sum(y_i - \hat{y}_i)^2}{n} \qquad (1.6)$$

Where:

y_i is the i value seen.

$\hat{y}_i$ correspond to the expected value

n = the number of occurrences

Smaller MSE implies better model prediction.

MSE was computed using R-software for both poisson and Negative Binomial models as shown in Table 1.1 below

Table 1.1: Mean squared error check for Poisson and Negative Binomial models

Model	**MSE**
Poisson	1.45
Negative Binomial	0.92

From Table 1.1 Negative Binomial model had lower MSE of 0.92 compared to Poisson model's 1.45, therefore Negative Binomial model provides better predictive accuracy than Poisson.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This chapter reviews theoretical and empirical literature on how mortality shocks influence fertility, focusing on the COVID-19 pandemic. It synthesizes global and regional findings, outlines the main explanatory mechanisms, and identifies methodological approaches previously applied in modeling fertility dynamics. The chapter also provides a detailed rationale for adopting the Negative Binomial Regression (NBR) model, supported by statistical justification using the dispersion test and Akaike Information Criterion (AIC). Finally, the review highlights the key knowledge gaps that this study seeks to address in the Kenyan context.

2.2 Theoretical perspectives on the mortality-fertility relationship

Mortality shocks have long been recognized as critical demographic events that influence fertility through multiple, sometimes opposing, mechanisms. The replacement hypothesis suggests that parents consciously or unconsciously have additional children to replace those lost [32, 3, 35, 20]. The insurance hypothesis posits that couples increase fertility preemptively when child mortality risk is perceived as high [26] . In contrast, tempo or postponement effects occur when mortality crises coincide with economic or psychosocial instability, leading to delayed childbearing [39]. Furthermore, economic-security models argue that pandemics and recessions typically depress fertility due to income loss, job insecurity, and uncertainty [1]. In the context of COVID-19, these mechanisms may interact: while replacement and insurance effects could raise fertility, pandemic-related disruptions to healthcare,

contraception, and employment may suppress it. The net impact thus depends on the magnitude of mortality, economic resilience, and access to reproductive health services.

2.3 Empirical evidence from past mortality shocks

Historical pandemics offer insight into the dynamic link between mortality and fertility. Studies of the 1918 influenza pandemic found a pronounced fertility decline roughly nine months after mortality peaks, attributed to both behavioral postponement and biological effects [26]. McBain [30] opines that during the 2014 peak of the Ebola epidemic in West Africa which killed fifty percent of infected individuals, family planning distribution declined tremendously in Liberia, Guinea and Sierra Leone. There was a baby boom after the epidemic in Liberia that was attributed to a rise in unintended pregnancies.Livi [26] examined how human fertility responded to mortality shocks as a result of famines, epidemics and natural calamities from past cases. The researcher concluded that there was evidence of a fairly common pattern. Fertility was seen to rebound after the crisis. These findings underscore that mortality shocks often produce short-term fertility deficits followed by recovery, though the magnitude varies with socioeconomic development, health infrastructure, and cultural norms. The historical evidence supports including lagged effects in models that link COVID-19 mortality to fertility outcomes.

2.4 COVID-19 and fertility

Globally, most studies document a temporary fertility decline following the onset of COVID-19, particularly in high income countries [1]. In middle- and low-income contexts, outcomes are more heterogeneous. Some report reduced fertility intentions due to financial instability [27] , whereas others note short-term increases linked to constrained contraceptive access [7]. In sub-Saharan Africa, the pan-

demic primarily affected fertility through interruptions in family planning services, mobility restrictions, and reduced incomes. In Kenya, early lockdowns disrupted contraceptive supply chains, leading to reports of increased unintended pregnancies among adolescents and low-income women[47]. Nevertheless, longitudinal surveys indicated that aggregate fertility intentions remained relatively stable [48], suggesting complex and context-specific effects. Despite these contributions, few studies explicitly link regional COVID-19 mortality shocks to realized fertility outcomes. Most existing analyses emphasize behavior or service access rather than demographic outcomes, underscoring the need for quantitative models that incorporate mortality directly.

2.5 Methodological approaches and rationale for the Negative Binomial Model

Fertility outcomes, such as the number of births per woman, are count variables. Early studies often employed Poisson regression, which assumes the equality of the mean and variance. However, real fertility data typically exhibit overdispersion (variance $>$ mean) due to unobserved heterogeneity differences in fecundity, marital stability, contraceptive access, and socio-cultural norms [10].

The Negative Binomial Regression (NBR) model introduces an additional dispersion parameter that accounts for extra Poisson variability, making it more flexible for demographic applications [20]. In addition to overdispersion, several other factors justify using NBR:

(i) Unobserved heterogeneity: Fertility varies across subgroups and regions, inflating variance.

(ii) Skewed count distributions: Fertility data often have many low counts and few very high counts.

(iii) Clustered data: Women nested within counties or households create correlated counts.

(iv) Zero inflation: NBR accommodates excess zeros better than Poisson.

To statistically justify the model, this study implemented both the dispersion test and the AIC comparison test:

Dispersion Test

Overdispersion was tested using the;

null hypothesis $H_0 : \theta = 0$ (Poisson adequacy)

against the;

alternative $H_1 : \theta > 0$

In R, the dispersion test() function from the AER package was used: The test was used to check whether Poisson assumption (Variance = Mean) holds

If P < 0.05: Reject null hypothesis.

This will imply that the data is overdispersed, so the Negative Binomial model is appropriate.

If P > 0.05: No significant overdispersion and therefore Poisson might be sufficient.

Table 2.1: The dispersion test results for Poisson model

Z	4.35
P-value	0.001

Since p < 0.05: Reject null hypothesis.

This implies that data is overdispersed, so the Negative Binomial model is appropriate.

AIC Test

The AIC provides a comparative measure of model fit, penalizing model complexity. Both Poisson and NBR models were estimated in R-software,and the output was as shown in Table 2.2 below

Based on Akaike Information Criterion, lower values immlies better fit

Table 2.2: AIC test results for Poisson and Negative Binomial models

Model	**df**	**AIC**
Poisson	5	1120.35
Negative Binomial	6	845.29

From Table 2.2, a lower AIC for the NBR model indicated superior goodness of fit. This finding was consistent with prior demographic applications where NBR outperformed Poisson when fertility counts were dispersed or heterogeneous [21]. Therefore, the choice of the Negative Binomial model in this study is justified both empirically (through dispersion and AIC tests) and theoretically (due to the nature of fertility data and unobserved heterogeneity).

2.6 Developed Studies using Poisson and Negative Binomial Regression Model

Modelling count data is a common task in sciences. Hurdle and Zero-inflated model are capable of handling over-dispersion and excess zeros. Notwithstanding different models used in modelling count data, Negative Binomial Regression has been reported as the main method for count data modelling [10, 24], especially when their is overdispersion.

The Poisson regression model has long been the standard approach for analyzing count data such as fertility, mortality, or morbidity events. It assumes that the mean and variance of the dependent variable are equal-an assumption often violated

in real demographic data [10]. Numerous fertility studies initially employed the Poisson model to estimate determinants of the number of children ever born or births within a given period. For instance, Adewale and Adepoju [2] modeled fertility behavior among Nigerian women using Poisson regression, identifying key socioeconomic predictors but noting unexplained variability in the data. Similarly, [29] applied a Poisson framework to child mortality counts in Malawi.

Even though the Poisson model gives simpler computational and interpretable coefficients, it has a major limitation when used to model fertility data. most of the empirical studies have oftenly reported over dispersion, where the variance of the number of births substantially exceeds the mean, this automatically violates the Poisson assumptions. This results to underestimated standard errors and inflated significance levels, hence compromising inference accuracy [21]. Not forgetting, the Poisson model assumes independence of events, which is unrealistic in fertility studies where unobserved heterogeneity, clustering (by household or region), and correlated behaviors are common [10].

These demerits have resulted into a methodological shift toward more flexible count models. The Negative Binomial Regression (NBR) model incorporates a dispersion parameter that captures unobserved heterogeneity and corrects for over dispersion, improving goodness of fit statistics such as the Akaike Information Criterion (AIC) and deviance measures. Comparative studies [21] consistently demonstrate that the NBR outperforms Poisson when analyzing fertility or mortality counts with heterogeneous populations.

Several demographic and public health studies have applied the Negative Binomial Regression (NBR) model to analyze count outcomes exhibiting overdispersion and heterogeneity. For instance, [17] employed the NBR model to identify determinants of fertility in Ethiopia and found it superior to the Poisson model based on AIC

and deviance statistics. Similarly, [29] applied NBR to model under-five mortality counts in Malawi, emphasizing its robustness to extra-Poisson variation. In health-related research, [21, 10] demonstrate that NBR consistently provides more reliable estimates where data variance exceeds the mean.

Consequently, while the Poisson model remains foundational in demographic research, its restrictive assumptions create gaps that justify the adoption of the Negative Binomial model in this study to examine fertility changes following COVID-19 mortality shocks in Kenya.

Choi [11] did a study using Poisson Regression and Negative Binomial on fertility of women among Mexican immigrants in the US. The aim of the study was to compare the study results with the official reports so as to help in establishing whether Poisson or Negative Binomial Regression was a reliable model. The researcher used data from the Mexican Life Survey the 2006-2010 NSFG and the 2002 NSFG. The respondents were aged between 15 to 49 years. The respondents were further grouped into 5 year age brackets. The ASFR was predicted using Poisson and Negative Binomial regression and it was established that NBR yielded results that were consistent with the official reports, Poisson model yielded results that were significantly different from the official reports.Therefore NBR was found to be a better model on modelling fertility.

Cameron and Trivedi [10] applied Poisson and Negative Binomial Models for count data which had prominence on the underlying count process and links to dual data on durations. But never applied it in modelling mortality shock to ascertain its effect on fertility. Both Negative Binomial model and Poisson Regression gave consistent outcomes.The major difference being that the Negative Binomial distribution was overdispersed,that is its variance is greater than its mean while Poisson had its variance equal to its mean.

Tsegaye [41] studied the Socio economic factors that influence women's fertility in Zambia in comparison with various other African countries. He employed the Negative Binomial model. The study focused on the total number of children ever born. The data was analysed alongside other variables that affect women including; age, education, marriage age, occupation etc. Poisson regression analysis and Negative Binomial analysis were used to confirm the hypotheses. NBR had a lower AIC value compared to Poisson model. It was found that all the factors influenced fertility, though in different degree

Ahmed and Ali [4] conducted a study on modeling the factors influencing fertility in Sudan. The models used in their study were Poisson Regression and Negative Binomial Regression. Secondary data from Sudan Household Health Survey (SHHS) (2010) was used. The results from this study indicated that both Poisson and Negative Binomial were viable to be used to predict and or analyse count data. Both models were able to predict the average number of children born by Sudanese Women. The findings also revealed that a significant relationship existed between fertility and various factors including age, gap between births, infant mortality etc. NBR had lower AIC, and more reliable standard errors than Poisson.

Amsler & Schmidt [5] used a Negative Binomial Regression model to analyze the number of bankruptcies in the banking sector, addressing the issue of overdispersion and providing more accurate risk assessments.

Mullahy [31] Applied the Negative Binomial distribution in econometrics, showing its utility in modeling overdispersed count data in the context of demand estimation. [31] argues that this distribution is often more appropriate than the Poisson model when count data exhibit overdispersion.

Cameron and Trivedi [10] In their seminal work on count data analysis, they applied

the Negative Binomial model to epidemiological data to model the incidence of diseases such as tuberculosis and HIV, showing how it addresses overdispersion better than the Poisson model.

Ghosh & Walker [18] Used Negative Binomial Regression to model the count of malaria incidences across different regions, showing the importance of overdispersion in understanding disease dynamics.

Virginia [44] modelled the effect of the great Ukrainian famine in 1933 mortality shock on fertility and human rate of aging. He used a generalized linear model with Poisson log-link function for the analysis. He investigated the impact of the great famine that had occurred in the year 1933 compared to the previous and the following years, i.e two models. An anova test was used to validate the models. His findings were that mortality shock led to an increase in fertility afterwards. The mortality curves during the years of the crisis, showed convergence at old ages.

Clement [12] did a study labeled "An analysis of fertility before, during and after the Rwanda genocide that had taken place in the year 1994". He noted that the genocide resulted in high mortality shock, this had an effect on fertility, having experienced the death of a child during the genocide increased both the need of having a child in the five years following the genocide and the total number of post-genocide births. Therefore there was strong evidence of a child replacement effect [40].

These applications underscore that fertility and health outcomes, which are discrete and often overdispersed, are better analyzed with NBR than with traditional Poisson models. The present study extends this methodological tradition by applying the NBR model to assess how COVID-19 mortality shocks influenced fertility rates in Kenya-a context where regional heterogeneity, clustered data, and unobserved

behavioral factors further justify the model's use.

The reviewed literature demonstrates that mortality shocks can produce diverse fertility responses, but empirical quantification in sub-Saharan Africa remains scarce. Historical and contemporary studies concur that the magnitude and direction of the fertility response depends on contextual factors and demographic structures. Not withstanding, existing analyses are limited by methodological simplicity and failure to integrate mortality indicators

By applying Negative Binomial Regression model, validated by the dispersion test and AIC tests, this study provides a statistically significant assessment of COVID-19's mortality - fertility relationship in Kenya. It thus extends both theoretical understanding and methodological practice in demographic modeling.

CHAPTER THREE

MODEL DESCRIPTION

3.1 Introduction

This chapter outlines the methodological framework adopted in the study. It presents the model assumptions, statistical formulation of the Negative Binomial Regression (NBR) model used to analyse fertility rates in Kenya, following the COVID-19 mortality shock. The model is developed from first principles as a Poisson Gamma mixture, from which the probability mass function (PMF), mean and variance, and log-likelihood are derived. Finally, the estimation procedure is discussed.

3.2 Model Assumptions

1. Observations Y_i are conditionally independent given covariates.
2. The log-link function $\log(\mu_i) = \mathbf{x}_i^\top \boldsymbol{\beta}$ is correctly specified.
3. The dispersion parameter α is constant across observations.
4. Overdispersion exists; if excessive zeros remain, zero-inflated variants may be tested using AIC or Vuong diagnostics.

3.3 The Negative Binomial Regression Model

Let Y_i denote the observed count (number of births) for the i-th woman, and let λ_i be the underlying Poisson rate. Conditional on λ_i,

$$Y_i \mid \lambda_i \sim \text{Poisson}(\lambda_i), \qquad P(Y_i = y_i \mid \lambda_i) = \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!},\ y_i = 0, 1, 2, \ldots \qquad (3.1)$$

Assume the random rate follows a Gamma distribution with shape $\alpha > 0$ and scale $\beta > 0$,

$$\lambda_i \sim \text{Gamma}(\alpha, \beta), \qquad f(\lambda_i) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \lambda_i^{\alpha-1} e^{-\lambda_i/\beta}, \tag{3.2}$$

with mean $E[\lambda_i] = \alpha\beta$ and variance $\text{Var}(\lambda_i) = \alpha\beta^2$.

The marginal distribution of Y_i is obtained by integrating the Poisson likelihood over the Gamma density:

$$P(Y_i = y_i) = \int_0^\infty P(Y_i = y_i \mid \lambda_i) f(\lambda_i)\, d\lambda_i$$

$$= \frac{1}{y_i!\,\Gamma(\alpha)\beta^\alpha} \int_0^\infty \lambda_i^{y_i+\alpha-1} \exp\big[-\lambda_i(1 + 1/\beta)\big]\, d\lambda_i. \tag{3.3}$$

Evaluating the integral using the Gamma identity $\int_0^\infty x^{k-1}e^{-cx}\,dx = \Gamma(k)/c^k$ gives

$$P(Y_i = y_i) = \frac{\Gamma(y_i + \alpha)}{\Gamma(\alpha) y_i!} \frac{1}{\beta^\alpha} \left(1 + \frac{1}{\beta}\right)^{-(y_i+\alpha)}. \tag{3.4}$$

Let $\mu_i = E[Y_i] = \alpha\beta$ so that $\beta = \mu_i/\alpha$ and $1 + 1/\beta = (\alpha + \mu_i)/\mu_i$. Substituting yields the standard NB form:

$$P(Y_i = y_i) = \frac{\Gamma(y_i + \alpha)}{\Gamma(\alpha)\, y_i!} \left(\frac{\alpha}{\alpha + \mu_i}\right)^\alpha \left(\frac{\mu_i}{\alpha + \mu_i}\right)^{y_i}, \qquad y_i = 0, 1, 2, \ldots \tag{3.5}$$

where $\mu_i > 0$ is the mean and $\alpha > 0$ is the dispersion (size) parameter.

3.4 Mean and Variance of the Model

From the Poisson–Gamma hierarchy:

$$E[Y_i] = E_{\lambda_i}[E(Y_i \mid \lambda_i)] = E[\lambda_i] = \mu_i, \tag{3.6}$$

$$\text{Var}(Y_i) = E_{\lambda_i}[\text{Var}(Y_i \mid \lambda_i)] + \text{Var}_{\lambda_i}[E(Y_i \mid \lambda_i)$$

$$= E[\lambda_i] + \text{Var}(\lambda_i) = \mu_i + \frac{\mu_i^2}{\alpha}. \tag{3.7}$$

$$\boxed{\text{Var}(Y_i) = \mu_i + \frac{\mu_i^2}{\alpha}} \tag{3.8}$$

When $\alpha \to \infty$, $\mathrm{Var}(Y_i) \to \mu_i$ and the distribution approaches the Poisson, confirming that the NB generalises the Poisson by allowing overdispersion.

3.5 Regression Specification

To relate the mean to covariates, a log link is used:

$$\log(\mu_i) = \mathbf{x}_i^\top \boldsymbol{\beta}, \qquad \mu_i = \exp(\mathbf{x}_i^\top \boldsymbol{\beta}), \tag{3.9}$$

where $\mathbf{x}_i$ is the vector of explanatory variables and $\boldsymbol{\beta}$ the vector of coefficients.

3.6 Log-Likelihood Function

Assuming independence, the joint likelihood is

$$L(\boldsymbol{\beta}, \alpha) = \prod_{i=1}^{n} \frac{\Gamma(y_i + \alpha)}{\Gamma(\alpha) y_i!} \left(\frac{\alpha}{\alpha + \mu_i}\right)^{\alpha} .10 \left(\frac{\mu_i}{\alpha + \mu_i}\right)^{y_i}. \tag{3.10}$$

Taking logarithms yields the log-likelihood:

$$\begin{aligned} \ell(\boldsymbol{\beta}, \alpha) &= \sum_{i=1}^{n} \Big\{ \ln \Gamma(y_i + \alpha) - \ln \Gamma(\alpha) - \ln(y_i!) + \alpha \ln\Big(\frac{\alpha}{\alpha + \mu_i}\Big) + y_i \ln\Big(\frac{\mu_i}{\alpha + \mu_i}\Big) \Big\} \\ &= \sum_{i=1}^{n} \Big[\ln \Gamma(y_i + \alpha) - \ln \Gamma(\alpha) - \ln(y_i!) + \alpha \ln(\alpha) - \alpha \ln(\alpha + \mu_i) + y_i \ln(\mu_i) - y_i \ln(\alpha + \mu_i) \end{aligned} \tag{3.11}$$

Note that the dispersion parameter appears as α,

3.7 Parameter Estimation

The maximum-likelihood estimates $(\hat{\boldsymbol{\beta}}, \hat{\alpha})$ are obtained by solving the score equations 3.12 and 3.13

$$\frac{\partial \ell}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{n} \left(y_i - \mu_i \frac{y_i + \alpha}{\alpha + \mu_i} \right) \mathbf{x}_i = \mathbf{0}, \tag{3.12}$$

$$\frac{\partial \ell}{\partial \alpha} = \sum_{i=1}^{n} \left[\psi(y_i + \alpha) - \psi(\alpha) + \ln\Big(\frac{\alpha}{\alpha + \mu_i}\Big) - \frac{\mu_i}{\alpha + \mu_i} \right] = 0, \tag{3.13}$$

where $\psi(\cdot)$ is the digamma function. Because these equations have no closed-form solutions, numerical algorithms such as Newton-Raphson or Fisher scoring are used.

Regression Coefficients $\beta_1, ..., \beta_k$, are estimated using the numerical optimization methods described above. After convergence, the coefficients will describe the relationship between the predictors $X_1, ..., X_p$ and the expected count.
The coefficients $\beta_1, ..., \beta_k$ represent the change in the log of the expected count μ_i for a one-unit change in the respective predictor variable, holding all other variables constant.

The Dispersion Parameter α is also estimated during the optimization process.
If α is close to 1, this suggests that the model is close to a Poisson regression model.
If $\alpha > 1$, the data is highly overdispersed, and the Negative Binomial model becomes better than a Poisson model.
In practice, software implementations maximise 3.13 directly.

In this study, five key predictors were used to model fertility:,i.e

(i) Marital status(M).

Under Marital status, the number of married women was considered, M is used to represent the same.

(ii) Induced abortion(A)

Under induced abortion, the number of women who had a pregnancy loss was considered, A is used to represent the same

(iii) Contraceptive use(C).

Under Contraceptive use, the number of women using Contraceptive was considered, C is used to represent the same

(iv) Sterility(S).

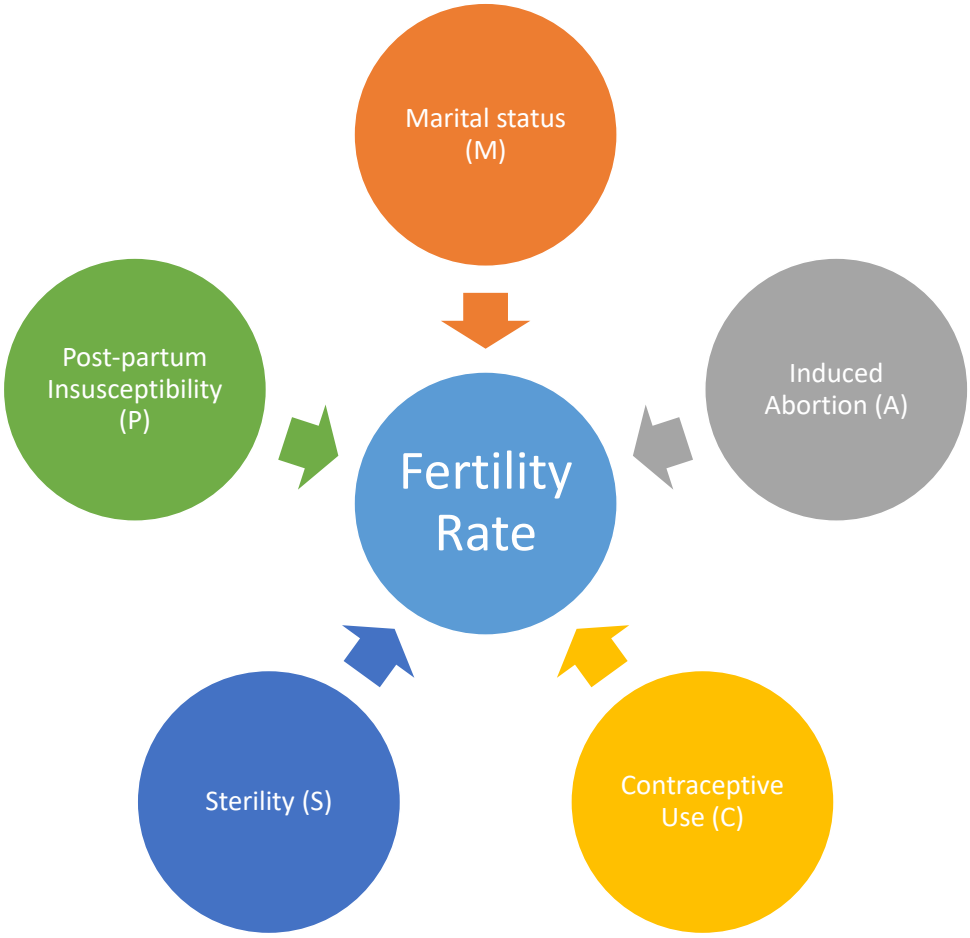
Under Sterility, the number of women who are sterile was considered, S is used to represent the same

(v) Post-partum insusceptibility(P).

Under Post-partum insusceptibility, the number of women who were insusceptible was considered, P is used to represent the same

This variables of interest are represented in Figure 3.1 below.

Figure 3.1: Visual display of fertility determinants of interest



These variables were selected based on their theoretical and empirical relevance to fertility behavior in Kenya, both before and after the COVID-19 pandemic.

Incorporating these predictors in the linear form of 3.9 the Poisson regression model becomes:

$$\ln(\mu) = \beta_0 + \beta_1 M + \beta_2 A + \beta_3 C + \beta_4 S + \beta_5 P \tag{3.14}$$

Taking the exponent of Equation (3.14), we obtain:

$$\mu = e^{\beta_0} \cdot e^{\beta_1 M} \cdot e^{\beta_2 A} \cdot e^{\beta_3 C} \cdot e^{\beta_4 S} \cdot e^{\beta_5 P} \tag{3.15}$$

Or equivalently:

$$\mu = \exp\left(\beta_0 + \beta_1 M + \beta_2 A + \beta_3 C + \beta_4 S + \beta_5 P\right) \tag{3.16}$$

Where:

- μ is the expected fertility rate,
- M, A, C, S, P represent marital status, induced abortion, contraceptive use, sterility, and post-partum insusceptibility, respectively,
- β_0 is the intercept,
- $\beta_1, \beta_2, \ldots, \beta_5$ are the regression coefficients.

To differentiate between the pre- and post-COVID-19 periods, the expected values are denoted as μ_b (before COVID-19) and μ_a (after COVID-19).

Before COVID-19:

$$\ln(\mu_b) = \beta_0 + \beta_1 M_b + \beta_2 A_b + \beta_3 C_b + \beta_4 S_b + \beta_5 P_b \tag{3.17}$$

$$\mu_b = \exp\left(\beta_0 + \beta_1 M_b + \beta_2 A_b + \beta_3 C_b + \beta_4 S_b + \beta_5 P_b\right) \tag{3.18}$$

After COVID-19:

$$\ln(\mu_a) = \beta_0 + \beta_1 M_a + \beta_2 A_a + \beta_3 C_a + \beta_4 S_a + \beta_5 P_a \tag{3.19}$$

$$\mu_a = \exp\left(\beta_0 + \beta_1 M_a + \beta_2 A_a + \beta_3 C_a + \beta_4 S_a + \beta_5 P_a\right) \tag{3.20}$$

After fitting the model, the coefficients are interpreted as follows:

- β_0: Represents the log of the expected fertility rate when all predictors are zero.
- $\beta_1, \ldots, \beta_5$: Represent the log-linear effects of the corresponding predictors on fertility. The exponential of each coefficient, e^{β_k}, gives the Incidence Rate Ratio (IRR), which quantifies the multiplicative change in the expected number of births per unit increase in the predictor.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 Introduction

This chapter presents and discusses the empirical results obtained from the Negative Binomial Regression (NBR) model used to examine the effect of the COVID-19 mortality shock on fertility in Kenya. The analysis proceeds in four stages. First, fitting the Negative Binomial Regression model to COVID-19 mortality shock free data set and to data set after COVID-19 is done and the parameters estimation Second, the regression results are presented and interpreted, the average fertility rates before and after the COVID-19 period are compared and related to the study hypothesis. Finally, the findings are discussed in light of existing demographic theory and prior empirical studies in sub-Saharan Africa.

The results are based on cross-sectional count data derived from Kenya Demographic and Health Survey microdata, partitioned into pre-COVID and post-COVID periods. All estimations were conducted using the NBR model as implemented in `R`, ensuring robustness to overdispersion and unobserved heterogeneity.

4.2 Fitting the Negative Binomial Regression model to COVID-19 mortality shock free data set in Kenya and Parameter Estimation

To fit a Negative Binomial regression model to compare the predictors before COVID-19 mortality shock, the study used the KDHS data of 2018. Five predictors of significant influence on fertility were extracted from the said KDHS survey of 2018. The main objective was to understand how these predictors (Marital status, Induced abortion, Contraceptive use, Sterility, and Post-partum insusceptibility)

influenced fertility rate before the COVID-19 shock.In that KDHS survey of 2018, 31,079 women were interviewed. The report had a number of explanatory variables, for this work predictors of significant inluence on fertility were extracted out.They included; the number of married women (M) under marital status, the number of women who lost pregnancy (A) under Induced abortion, the number of women using Contraceptive (C) under Contraceptive use, the number of women who are sterile(S)under Sterility and the number of women who were insusceptible (P) under Post-partum insusceptibility.

Table 4.1: The Table of Extracted explanatory variables before COVID-19

Predictors	**Numbers**	**Percentage**
Marital status (M)	18,549	59.7
Induced abortion (A)	2,176	7.0
Contraceptive use (C)	13,240	42.6
Sterility (S)	1,150	3.7
Post-partum insusceptibility (P)	746	2.4

The study used Negative Binomial regression to model fertility rates before the COVID-19 shock. The dependent variable was the number of births, and the independent variables were the predictors; Marital status, Induced abortion, Contraceptive use, Sterility, and Post-partum insusceptibility.

The Negative Binomial regression model was fitted before COVID-19 as follows;

$$\ln(\mu_b) = 0.800 - 0.068M_b - 0.045A_b + 0.210C_b - 0.080S_b - 0.110P_b \tag{4.1}$$

Taking the exponential of Equation (4.1), the model becomes:

$$\mu_b = \exp(0.800 - 0.068M_b - 0.045A_b + 0.210C_b - 0.080S_b - 0.110P_b) \tag{4.2}$$

Where:

- μ_b is the expected number of births before COVID-19,

- M_b, A_b, C_b, S_b, and P_b represent marital status, induced abortion, contraceptive use, sterility, and post-partum insusceptibility, respectively.

To interpret the effect of each predictor, the coefficients are exponentiated to obtain the Incidence Rate Ratios (IRR), defined as:

$$\text{IRR} = e^{\beta}$$

The IRR indicates the multiplicative change in the expected number of births associated with a one-unit increase in the corresponding predictor, holding all other variables constant. An IRR greater than 1 suggests a positive association with fertility, while an IRR less than 1 indicates a negative association.

4.2.1 The results of the Model before COVID-19

The results of the Model fitting is presented as shown in Table 4.2.

Table 4.2: Table of the estimated parameters of the regression model before COVID-19

Parameter	**Estimate**	**Std. Error**	**Z-value**	**P-value**	**Exp(Estimate)**
Intercept	0.800	0.003	266.67	<2e-16	2.225
M_b	-0.068	0.015	8.00	<2e-16	0.934
A_b	-0.045	0.010	-4.50	0.0001	0.956
C_b	0.210	0.030	7.00	<2e-16	1.234
S_b	-0.080	0.020	-4.00	0.0017	0.923
P_b	-0.110	0.035	-3.14	<2e-16	0.896
α	=	1.537			

Table 4.2 presents the results of the fitted model conducted on data from before the COVID-19 pandemic. Each term represents a predictor in the model, and the associated estimate, standard error, Z-value, P-value, and exponentiated estimate (Exp(Estimate)) provide insights into the influence of these predictors on the outcome variable.

α represents the dispersion parameter,the value of 1.537 which is greater than one implies that Poisson assumption of mean and variance being equal doesn't hold. therefore giving more justification on the use of Negative Binomial Regression.

The intercept has an estimate of 0.800 with a very small standard error of 0.003. The extremely high z-value of 266.67 and a p-value of less than ($<$ 2e-16) indicate that this result is highly statistically significant. The exponentiated estimate (Exp(Estimate)) is 2.225, suggesting that the baseline count of the outcome variable is 2.225 times the reference level when all predictors are at zero.

Predictor M has an estimate of -0.068 and a standard error of 0.015. With a z-value of 8.00 and a p-value less than 2e-16, the result is highly significant. The Exp(Estimate) of 0.934 indicates that Predictor M is associated with a 6.6% decrease in the outcome count.

The estimate for Predictor A is -0.045 with a standard error of 0.010. The z-value of -4.50 and p-value of 0.0001 indicate a statistically significant negative association. The Exp(Estimate) of 0.956 suggests that Predictor A is associated with a 4.4% decrease in the outcome count.

Predictor C shows a positive estimate of 0.210 and a standard error of 0.030. The predictor has a Z-value of 7.00. The P-value for the same predictor is less than 2e-16, this indicates that the result is highly significant. The Exp(Estimate) of 1.234 indicates a 23.4% increase in the outcome count associated with Predictor C.

Predictor S has an estimate of -0.080 and a standard error of 0.020. The z-value of -4.00 and p-value of 0.0017 suggest a statistically significant negative association. The Exp(Estimate) of 0.923 means Predictor S is associated with a 7.7% decrease in the outcome count.

Predictor P shows an estimate of -0.110 with a standard error of 0.035. The z-value

of -3.14 and a highly significant p-value of less than 2e-16 indicate a strong negative association. The Exp(Estimate) of 0.896 suggests that Predictor P is associated with a 10.4% decrease in the outcome count.

Overall, all predictors in the model are statistically significant, indicating that each has a meaningful effect on fertility behavior before the pandemic. Among the predictors, contraceptive use shows the strongest positive influence, while Post-partum insusceptibility exerts the most pronounced negative effect.

4.3 Fitting the Negative Binomial Regression model to data set affected by COVID-19 mortality shock in Kenya and Parameter Estimation

To fit a Negative Binomial regression model to compare the predictors After COVID-19 mortality shock, the study used the KDHS data of 2022. Five predictors of significant influence on fertility were extracted from the said KDHS survey of 2022. The main objective was to understand how these predictors (Marital status, Induced abortion, Contraceptive use, Sterility, and Post-partum insusceptibility) influence the count of events (number of women interviewed) after the COVID-19 shock.In that KDHS survey of 2022 32,145 women were interviewed. the report had a number of explanatory variables, for this work predictors of significant influence on fertility just like for before COVID-19 were extracted out.They included; the number of married women (M) under marital status, the number of women who lost pregnancy (A) under Induced abortion, the number of women using Contraceptive (C) under Contraceptive use, the number of women who are sterile(S)under Sterility and the number of women who were insusceptible (P) under Post-partum insusceptibility.

Table 4.3: The Table of Extracted explanatory variables after COVID-19

Predictors	**Numbers**	**Percentage**
Marital status (M)	17,822	55.4
Induced abortion (A)	1,770	5.5
Contraceptive use (C)	18,702	58.2
Sterility (S)	804	2.5
Post-partum insusceptibility (P)	547	1.7

The study used Negative Binomial Regression to model fertility rate after the COVID-19 mortality shock. The dependent variable was the number of births, and the independent variables were the predictors; Marital status, Induced abortion, Contraceptive use, Sterility, and Post-partum insusceptibility.

The Negative Binomial regression model was fitted after COVID-19 as follows;

$$\ln(\mu_a) = 1.000 - 0.070M_a - 0.050A_a + 0.150C_a - 0.100S_a - 0.200P_a \tag{4.3}$$

Taking the exponential of both sides of 4.3 yields:

$$\mu_a = \exp(1.000 - 0.070M_a - 0.050A_a + 0.150C_a - 0.100S_a - 0.200P_a) \tag{4.4}$$

Where:

(i) μ_a is the expected fertility count after COVID-19.

(ii) M_a, A_a, C_a, S_a, and P_a represent Marital status, Induced abortion, Contraceptive use, Sterility, and Post-partum insusceptibility respectively, for the post-COVID-19 period.

(iii) The coefficient values represent the log change in the expected count of fertility for a one-unit change in the predictor.

To interpret the magnitude of each predictor's effect in terms of incidence rate ratios (IRR), we exponentiate the regression coefficients as shown in Equation 4.5:

$$\text{IRR} = e^{\beta} \tag{4.5}$$

The IRR reflects the factor by which the expected count of the outcome variable (fertility) changes for a one-unit increase in the predictor variable. An IRR greater than 1 implies an increase in fertility, while an IRR less than 1 implies a decrease.

4.3.1 The results of the Model after COVID-19

The results of the Model fitting is presented as shown in Table 4.4.

Table 4.4: Table of the Estimated parameters of the Regression model After COVID-19

Parameter	**Estimate**	**Std. Error**	**Z-value**	**P-value**	**Exp(Estimate)**
Intercept	1.000	0.001	1000.00	<2e-16	2.718
M_a	-0.07	0.020	-3.5	0.0005	0.933
A_a	-0.050	0.020	-2.50	0.012	0.951
C_a	0.150	0.025	6.00	<2e-16	1.162
S_a	-0.100	0.030	-3.33	0.0009	0.905
P_a	-0.200	0.040	-5.00	<2e-16	0.819
θ	=	1.572			

Table 4.4 present the results of the fitted model applied to post-COVID-19 fertility data. The dispersion parameter $\theta = 1.572$, being greater than one, confirms the presence of overdispersion in the data. This validates the choice of a Negative Binomial model over the Poisson regression, which assumes equidispersion (i.e., variance equal to the mean).

Intercept: The intercept is estimated at 1.000 with a negligible standard error (0.001), indicating high precision. The extremely high z-value (1000.0) and a p-value less than $2e^{-16}$ denote strong statistical significance. The exponentiated estimate, $e^{1.000} = 2.718$, implies that the baseline expected count of fertility (when all predictors are at reference level) is approximately 2.718.

Marital Status (M): Marital status has a coefficient of -0.070, with a standard error of 0.020 and a statistically significant P-value of 0.0005. The Incidence Rate Ratio (IRR) is $e^{-0.070} \approx 0.933$, this suggests that being married (or a one-unit increase in marital status variable) is associated with a 6.7% reduction in the expected fertility count. This may reflect delayed childbearing or disruption in family formation dynamics during the pandemic.

Induced Abortion (A): Induced Arbortin has a coefficient of -0.050, this indicates a negative association with fertility. The IRR of $e^{-0.050} \approx 0.951$ implies a 4.9% decrease in fertility for each unit increase in abortion incidence. This may be due to heightened emotional, economic, or service related constraints following the pandemic that led to increased pregnancy terminations.

Contraceptive Use (C): Contraceptive has a positive coefficient of 0.150. The IRR of $e^{0.150} \approx 1.162$ implies that contraceptive use is associated with a 16.2% increase in the expected count. This counter intuitive result could reflect greater reporting, improved access post-lock downs, or increased use among certain groups, possibly as a response to heightened economic uncertainty.

Sterility (S): The coefficient is -0.100, with an IRR of $e^{-0.100} \approx 0.905$, suggesting a 9.5% reduction in fertility associated with sterility. This is expected as biological inability to conceive directly reduces the likelihood of pregnancy outcomes.

Postpartum In susceptibility (P): The coefficient for Postpartum In susceptibility is -0.200, indicating the strongest negative association among all the predictors. The IRR of $e^{-0.200} \approx 0.819$ shows that women in a state of postpartum in susceptibility are 18.1% less likely to contribute to the fertility count, as they are temporarily not at risk of conception due to lactational amenorrhea or postpartum abstinence.

Overall, all predictors were statistically significant and showed meaningful effects on fertility rates after the COVID-19 pandemic. Notably, postpartum insusceptibility had the most substantial suppressive effect, while contraceptive use surprisingly showed a positive association—highlighting a complex behavioral and service-access response landscape in the aftermath of the pandemic.

4.4 Comparison of the result of the model on KDHS data before and after COVID-19

4.4.1 Density Curve of the Negative Binomial Regression model on KDHS datasets before and after COVID-19

Figure 4.1: Density Curve showing the behavior of fertilty before and after COVID-19

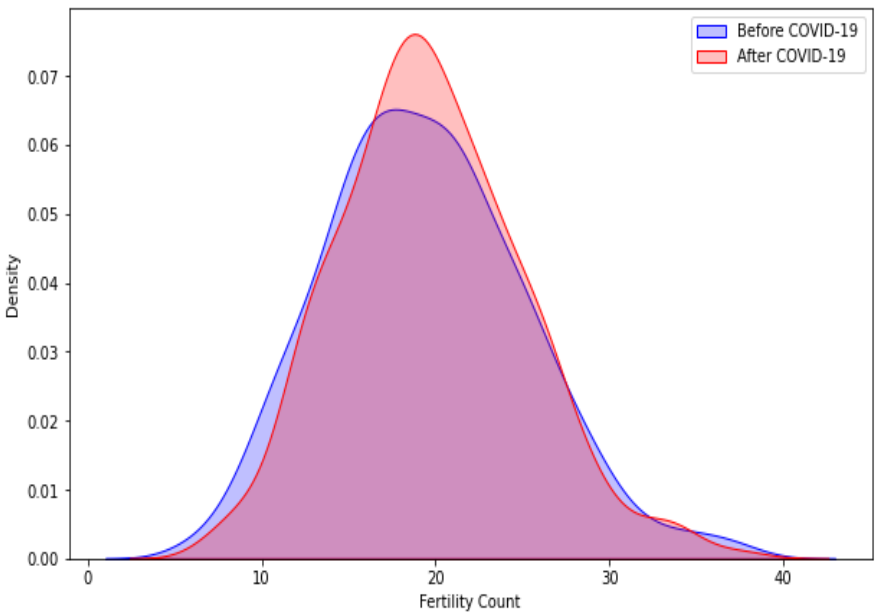


Figure 5.1 presents a density curve derived from the Negative Binomial Regression model applied to the KDHS datasets collected before and after the COVID-19 pandemic. This visual representation offers insight into the distribution and concentration of fertility rates across both time periods.

The curve for the after-COVID-19 period is more peaked and more peaked compared to the before-COVID-19 curve. This indicates that fertility rates not only increased after the pandemic but also became more concentrated around a central value. The heightened peak suggests reduced variability in fertility outcomes, implying a more uniform reproductive behavior among women during the post-pandemic period. Several factors may explain this observation, including a sense of economic recovery, improved healthcare access, and shifts in reproductive preferences or behavior.

Also, the density curve for the before-COVID-19 period is shifted to the left of the after-COVID-19 curve. This leftward shift indicates that fertility rates were generally lower before the pandemic, with a larger proportion of women reporting fewer births. In contrast, the post-pandemic curve is centered around higher fertility counts, indicating an upward shift in average fertility.

These trends could be attributed to various contextual factors such as:

(i) **Delayed childbearing :**

(ii) **Psychological and behavioral effects:**

(iii) **Changes in access to reproductive health services:**

(iv) **Perceived stability after the pandemic:**

In conclusion, the density curve supports the hypothesis that fertility increased and became more concentrated after COVID-19, reflecting behavioral, social, and service-related adjustments during the recovery phase.

4.4.2 Result from model fitting on KDHS data before and after COVID-19 compared

This section compares the fitted Negative Binomial Regression models based on KDHS data collected before and after the COVID-19 pandemic. The analysis draws upon Tables 4.2 (pre-pandemic) and 4.4 (post-pandemic) to interpret shifts in predictor effects.

Predictor M (Marital Status) : Before COVID-19, the coefficient estimate was -0.068 with an exponentiated value of $e^{-0.068} \approx 0.934$. After COVID-19, the coefficient was -0.070, with $e^{-0.070} \approx 0.933$. This indicates a consistent negative association between marital status and fertility both before and after the pandemic, with a slightly greater reduction in the post-COVID period.This also implies that before COVID -19,for one unit increase in M the expected count is multiplied by ≈ 0.912, While after COVID-19, for one unit increase in M the expected count is multiplied by ≈ 0.933 i.e for one unit increase in M reduces the expected count by 0.07

Predictor A (Induced Abortion) : Before the pandemic, the coefficient was -0.045 (IRR ≈ 0.956), and after the pandemic, it became -0.050 (IRR ≈ 0.951). This slight decline in IRR implies a marginally stronger negative effect of abortion on fertility post-COVID-19.This implies that before COVID -19,for one unit increase in A the expected count is multiplied by ≈ 0.956, While after COVID-19, for one unit increase in A the expected count is multiplied by ≈ 0.956

Predictor C (Contraceptive Use) : Pre-COVID, the coefficient was 0.210 (IRR ≈ 1.234), while post-COVID it dropped to 0.150 (IRR ≈ 1.162). Despite the drop, the variable maintained a positive influence on fertility, indicating contraceptive use became slightly less impactful in increasing fertility counts after the pandemic.This

implies that before COVID -19,for one unit increase in C the expected count is multiplied by ≈ 1.234, While after COVID-19, for one unit increase in M the expected count is multiplied by ≈ 1.162

Predictor S (Sterility) : The coefficient before COVID-19 was -0.080 (IRR ≈ 0.923), and after COVID-19 it became -0.100 (IRR ≈ 0.905).This implies that before COVID -19,for one unit increase in S the expected count is multiplied by ≈ 0.923, While after COVID-19, for one unit increase in S the expected count is multiplied by ≈ 0.905

Predictor P (Postpartum Insusceptibility) : The coefficient pre-pandemic was -0.110 (IRR ≈ 0.896), which decreased further to -0.200 post-pandemic (IRR ≈ 0.819). This reflects a significant increase in the negative impact of postpartum insusceptibility on fertility after the pandemic.This implies that before COVID -19,for one unit increase in P the expected count is multiplied by ≈ 0.896, While after COVID-19, for one unit increase in P the expected count is multiplied by ≈ 0.819

The cumulative effect of predictors on the expected fertility count is evaluated below for both periods:

4.4.3 Estimated Count Before COVID-19

For one unit increase in M_b, A_b, C_b, S_b and P_b, the expected count is multiplied by approximately 2.225(intercept), 0.912, 0.956, 1.234, 0.923 and 0.896.

$$\approx (1 \times 2.225) + (1 \times 0.912) + (1 \times 0.956) + (1 \times 1.234) + (1 \times 0.923) + (1 \times 0.896)$$

$$\mu_b \approx 2.225 + 0.934 + 0.956 + 1.234 + 0.923 + 0.896$$

$$\mu_b \approx 7.168 \tag{5.1}$$

4.4.4 Estimated Count After COVID-19

Similarly, for one unit increase in M_a, A_a, C_a, S_a andP_a, the expected count is multiplied by approximately 2.718(intercept), 0.933, 0.951, 1.162, 0.905 and 0.819.

$$\approx (1 \times 2.718) + (1 \times 0.933) + (1 \times 0.951) + (1 \times 1.162) + (1 \times 0.905) + (1 \times 0.819)$$

$$\mu_a \approx 2.718 + 0.933 + 0.951 + 1.162 + 0.905 + 0.819$$

$$\mu_a \approx 7.488 \tag{5.2}$$

4.5 Testing the Hypotheses of μ_a and μ_b

The mean fertility rates obtained from the Negative Binomial Regression models are as follows:

$\mu_a = 7.488$ (after COVID-19)

$\mu_b = 7.146$ (before COVID-19)

Hypotheses

- Null hypothesis (H_0): $\mu_a = \mu_b$ (no difference in means)
- Alternative hypothesis (H_1): $\mu_a \neq \mu_b$ (means are significantly different)

In the NBR model, the expected count is $\mu_i = \exp(X_i^\top \beta)$. Therefore,

$$\frac{\mu_a}{\mu_b} = \exp(\beta_{\text{covid}}).$$

The statistical significance of this coefficient directly tests whether $\mu_a = \mu_b$.

From R the $\beta_{\text{covid}} = 0.047$ with $p < 0.05$, the $\exp(0.047) = 1.048$, this imply that fertility after COVID-19 increased by approximately 4.8%.

Since the p-value for `covid_period` < 0.05, we reject H_0 and conclude that fertility changed significantly after COVID-19.

Therefore, $\mu_a = 7.488$ is significantly higher than $\mu_b = 7.146$,

Figure 4.2: Visual comparison of estimated fertility counts before and after COVID-19

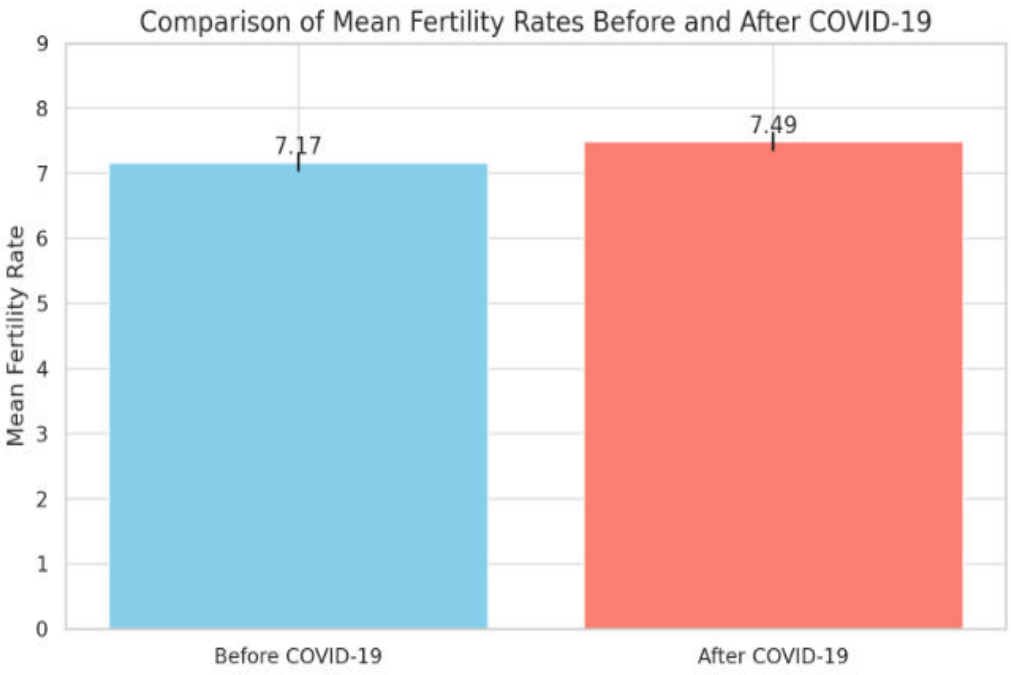


Figure 4.2 shows a bar plot comparing the mean fertility counts before and after the COVID-19 pandemic as estimated by the Negative Binomial Regression model. The plot reveals a modest increase in expected fertility count from approximately 7.15 before the pandemic to 7.49 after the pandemic. This visual plot supports the analytical comparison provided in Equations (5.1) and (5.2), reinforcing the finding that fertility behavior shifted in the wake of COVID-19.

The mean fertility rate increased from approximately 7.15 before the pandemic to 7.49 after.

The plot includes error bars to reflect the standard errors, indicating the precision of the estimates.

CHAPTER FIVE

CONCLUSION AND RECOMMENDATIONS

5.1 Introduction

This chapter summarises the key findings of the study and presents policy recommendations based on the results of the Negative Binomial Regression (NBR) analysis. The study investigated the effect of the COVID-19 mortality shock on fertility in Kenya, with specific emphasis on identifying whether fertility behaviour changed after the pandemic.

5.2 Summary of Findings Based on Study Objectives

The study employed the Negative Binomial Regression model to analyse fertility count data from the Kenya Demographic and Health Surveys. Negative Binomial Regression model was specified and fitted to two KDHS data sets, one before COVID-19 and the other after COVID-19 and their parameter estimated.Model adequacy test employed included dispersion test and AIC test. For AIC test Poisson model had a value 1120.35 wheres NBR had 845.29, thus NBR was found to be superior to Poisson because of lower AIC value(page 16). Dispersion test for Poisson (page 15) yielded P-value of 0.01 which was lower than 0.05 implying that the data was over dispersed. Test for better predictive Accuracy was done by MSE (page 11), for Poisson the result was 1.45 while for NBR the result was 0.92, this concluded that NBR had a better predictive accuracy than Poisson.Model adequacy tests confirmed that the NBR model outperformed the Poisson regression, effectively addressing the issue of overdispersion present in the fertility data. The model demonstrated that the mean fertility level after COVID-19 was $\mu_a = 7.488$

slightly higher than before COVID-19 which was $\mu_b = 7.146$. The two values were statistically significant at α =0.05 This small but statistically significant increase suggests that fertility responded positively to the mortality shock, consistent with compensatory or replacement fertility behaviour.

5.3 Conclusion

The findings of this study concluded that the COVID-19 mortality shock led to statistically significant increase in fertility in Kenya, mediated by behavioural and structural factors. By linking empirical findings to demographic theory, this study underscores the importance of integrating fertility and mortality policies within Kenya's broader health and development agenda.

5.4 Policy Recommendations

Findings from this study have several important implications for demographic and health policy in Kenya:

1. **Reproductive health surveillance and data systems to be strengthened .** There should be continuous monitoring of fertility and mortality trends which should be institutionalized to detect early demographic responses to health crises and guide adaptive policy interventions.

2. **Fertility and mortality policies should be integrated within a broader population resilience framework.** Policies should factor in the fact that mortality shocks can trigger behavioral fertility responses. Therefore integrated population health strategies are required to maintain balance between mortality reduction and sustainable fertility levels.

3. **Access to reproductive and family planning services chould be enhanced during crises.** The temporary disruptions in health services during COVID-19 probably contributed to fertility changes. Ensuring uninterrupted access to contraception and maternal care during emergencies is essential to avoid unintended fertility surges.

4. **Demographic modeling should be used to inform post crisis recovery planning.** more studies incorporating advanced models like the Negative Binomial Regression should guide policy on health recovery, social protection, and long-term demographic sustainability.

5.5 Areas for Further Research

Based on the findings of this study,several areas warrant further investigation to deepen our understanding between COVID 19 mortality shocks and fertility rates in Kenya

(1) Future research should examine whether the observed fertility decline during the pandemic which rebounds after is temporary or if it leads to sustained demographic shifts.

(2) A longitudinal analysis covering post pandemic years would help assess fertility recovery patterns.

REFERENCES

[1] Aassve, A., et al. (2021). Fertility and the COVID-19 pandemic: Early evidence from high-income countries. *Proceedings of the National Academy of Sciences*, 118(36), e2105709118.

[2] Adewale, A. S., et al. (2019). Determinants of fertility among women in Nigeria: Evidence from Poisson regression analysis.. *African Population Studies*, 33(2), 4512–4525.

[3] Agadjanian, V., & Prata, N. (2002). War, peace, and fertility in Angola. Demography.;39:215–231.

[4] Ahmed, H. M., & Ali, H. M. H. (2016). *Using count regression models to determine factors influencing fertility of Sudanese Women.*

[5] Amsler, C., & Schmidt, P. (2011). "A Negative Binomial Approach to Modelling Bankruptcy." Journal of Financial Econometrics, 9(4), 367-395.

[6] Athukorala, P.C., & Resosudarmo, B.P. (2000). The Indian Ocean tsunami: Economic impact, disaster management, and lessons. *Asian Economic Papers*, vol. 4, pp. 1—39.

[7] Bietsch, K., et al. (2021). Family planning during and after the COVID-19 pandemic: Evidence from sub-Saharan Africa. BMJ Global Health, 6(1), e004139.

[8] Bongaarts, R. & John, K. (2017). Africa's Unique Fertility Transition. *Population and Development Review*, vol (43), pp. 39-58.

[9] Bloom, et al. (2011). Implications of Population Aging for Economic Growth. *Oxford Review of Economic Policy*, vol 4(26), pp. 583-612.

[10] Cameron, A.C. & Trivedi, P.K. (2013). *Regression analysis of count data*, vol. 30. Cambridge Univ Pr., Handbook Stat. 14, pp 363–391.

[11] Choi, K. (2014). Fertility in the context of Mexican Migration to the United States: A case for incorporating the pre migration fertility of immigrants. *Demographic Research*, vol 130(24), pp. 703-737.

[12] Clement, B. (2006). Fertility in Rwanda: Impact of Genocide. "An analysis of fertility before, during and after the 1994 genocide". Project Submitted in Partial Fulfillment of the Requirements for the MSc Degree of Science, in the Department of Statistics, University of the Western Cape.

[13] Crawley, M. J. (2012). The R book, John Wiley & Sons.

[14] Davis, K. & Blake, J. (2017). *Social Structure and Fertility: An Analytical Framework. Economic Development and Cultural Change*, vol. 4(4), pp. 211-253.

[15] Dick, R., et al. (2007). "National Collegiate Athletic Association injury surveillance system commentaries:" introduction and methods. Journal of athletic training 42(2): 173.

[16] Doocy, S., Gorokhovich, Y., Burnham, G., Balk, D., & Robinson, C. (2007). *Tsunami mortality estimates and vulnerability mapping in Aceh, Indonesia. American Journal of Public Health*. 2007;97(Suppl 1):S146–S151.

[17] Gebre, G. G., Tadesse, M., & Teklu, B. (2021). Modelling determinants of fertility using count regression models: Evidence from Ethiopia. Journal of Population Studies, 28(3), 45–59.

[18] Ghosh, S., & Walker, M. (2009). "Modeling Malaria Incidence with Negative Binomial Regression." American Journal of Epidemiology, 169(7), 882-891.

[19] Gourieroux, C. (2000). *The econometrics of discrete positive variables:"The Poisson model". Econometrics of Qualitative Dependent variables*. York: Cambridge University Press.pp.270-83

[20] Heuveline, P., & Poch, B. (2007). The Phoenix population: Demographic crisis and rebound in Cambodia. *Demography*, vol 44, pp. 405–426.

[21] Hilbe, J. M. (2011). Negative binomial regression. Cambridge University Press.

[22] Islam, H.,*et al.* (2020). A Generalized overview of SARS-CoV-2: where does the current knowledge stand? *Electron J Gen Med.*, vol 17:em251.

[23] Jester, B., Uyeki, T., & Jernigan, D. (2018). Readiness for responding to a severe pandemic 100 years after 1918. *Am J Epidemiol.*

[24] Kleiber, C., Zeileis, A.,& Jackman, S. (2008). Regression models for count data in R. *J. Stat. Softw.* 27 (8) 1–25 .

[25] Lee, C.H., Huang, N., Chang, H.J., Hsu, Y.J., Wang, M.C., & Chou, Y.J. (2005) . The immediate effects of the severe acute respiratory syndrome (SARS) epidemic on childbirth in Taiwan. *BMC Public Health*, vol. 5, pp. 30.

[26] Livi,B.M. (2001). Demographic shocks: The view from history ." in Seismic shifts: The Economic Impact of Demographic Change.

[27] Luppi, F., Arpino, B., & Rosina, A. (2020). The impact of COVID-19 on fertility plans in Italy, Germany, France, Spain, and the UK. Demographic Research, 43(1), 1399–1412.

[28] Mamelund, S.E. (2012). Fertility fluctuations in times of war and pandemic influenza. *J Infect Dis.* vol 206, pp. 140—141.

[29] Manda, S. O. M., et al. (1999). A Poisson regression model of child mortality in Malawi. Journal of Biosocial Science, 31(2), 217–231.

[30] McBain., *et al.* (2016) https://www.thelancet.com/action/showPdf?pii=S0140-6736%2816%2931895-5

[31] Mullahy, J. (1986). Specification and testing of some modified count data models. Journal of Econometrics, 33(3), 241-258.

[32] Nobles, J., Frankenberg, E., & Thomas, D. (2015) The effects of mortality on fertility: population dynamics after a natural disaster. *Demography*, vol 52, pp. 15–38.

[33] Pfefferbaum, B., & North, C.S. (2020). Mental health and the COVID-19 pandemic. *N Engl J Med.*, vol 383, pp. 510–512.

[34] Pitkanen, K.J.(1992). The patterns of mortality during the Great Finnish Famine in the 1860s, *Acta Demographica*, pp. 81-102.

[35] Rodgers, J.L.,John. C.A., & Coleman, R.(2005). Did fertility go up after the Oklahoma City bombing? An analysis of births in metropolitan counties in Oklahoma. *Demography*, vol. 42, pp. 675–692.

[36] Ronald, & Lee, C. (2016). Macroeconomics, Demographic Transition, and Economic Growth. *Journal of Economic Perspectives*, vol 4 (30), pp. 257-280.

[37] Rossi, R.J. (2018). *Mathematical Statistics : An Introduction to Likelihood Based Inference*. New York: John Wiley and Sons. p. 227. ISBN 978-1-118-77104-4.

[38] Santos, C. & Barret, C.B. (2011). Social network capital,economic mobility and poverty traps. *J Econ Inequal*. doi: 10.1007/s10888-011-9164-5

[39] Sobotka, T., Jasilioniene, A., Galarza, A. A., et al. (2021). Baby bust in the wake of the COVID-19 pandemic?. *Human Reproduction*, vol 4(3), 36(10), 2896–2910.

[40] Tilman, B., Kati, K., Michele, D.M., & Roberto, N. (2019) The Effects of Conflict on Fertility: Evidence From the Genocide in Rwanda; doi: 10.1007/s13524-019-00780-8

[41] Tsegaye, T. (2010). Cross sectional study of women Employment and Fertility in Ethiopia. Master's thesis ,Department of Economic History, Lund University.

[42] Turpeinen, O. (1979). Fertility and mortality in Finland since 1750. *Popul Stud.*, vol 33, pp. 101–114.

[43] Upadhyay, H.P.,& Bhandari, K.R. (2017). Factors associated with children ever born: a case study of somadi village development committee of Palpa District, Nepal, *Adv. J. Soc. Sci*, vol 1(1), pp. 15-–29.

[44] Virginia, Z. (2012). "Mortality shocks and the human rate of aging, " MPIDR working Papers WP-2012-019, Max Planck Institute for Demographic Research.

[45] WHO, Ebola Response Team. (2014). Ebola virus disease in West Africa—the first 9 months of the epidemic and forward projections, *N Engl J Med. 371:1481–95*. doi: 10.1056/NEJMoa1411100.

[46] WHO.(2020). "Corona virus disease outbreak," World Health Organization. *https://www.who.int/news-room/detail/27-04-2020-who-covid-19*.

[47] Wood, S. N., et al. (2021). Contraceptive dynamics during COVID-19 in Kenya: Service disruptions and recovery. BMJ Global Health, 6(8), e005925.

[48] Winkelmann, R. & Zimmermann, K.F. (1994). Count data models for demographic data. *Mathematical Population studies*, vol 4(3), pp. 205-221.

APPENDICES

Appendix : R Code for Model Estimation

R Procedure for Fitting Negative Binomial Regression Model

Required packages library(readxl)

library(TSstudio)

library(lmtest)

library(MASS); For glm.nb()

library(pscl); For dispersion tests

library(AER) ; For dispersiontest()

library(ggplot2); For visualizatio

Load datasets (replace file paths)

$data_before <- -read.csv("KDHS2018.csv")$

$data_after <- -read.csv("KDHS2022.csv")$

Model 1: Before COVID-19

$model_before <- -glm.nb(Fertility\ Marital+Abortion+Contraceptive+Sterility+$

$Postpartum, data = data_before)$

$summary(model_before)$

Model 2: After COVID-19

$model_after <- -glm.nb(Fertility\ Marital+Abortion+Contraceptive+Sterility+$

$Postpartum, data = data_after)$

$summary(model_after)$

Dispersion test for Poisson adequacy

$pois_model <- -glm(Fertility\ \ Marital + Abortion + Contraceptive + Sterility +$

$Postpartum, family = poisson, data = data_before)$

$dispersiontest(pois_model)$

Compare AIC values

$AIC(model_before, model_after, pois_model)$

Compute Incidence Rate Ratios (IRR)

$exp(coef(model_before))$

$exp(coef(model_after))$

Predicted counts and density visualization

data-before $pred <- predict(model_before, type = "response")$

data-after $pred <- predict(model_after, type = "response")$

$ggplot() + geom_density(aes(data_beforepred), fill = "skyblue", alpha = 0.5) +$

geom-density(aes(data-after$pred), fill = "orange", alpha = 0.5) + labs(title =$

$"FertilityCountDensityBeforeandAfterCOVID-19", x = "PredictedFertilityCount", y =$

$"Density") + theme_minimal()$

Mean Squared Error (MSE)

$mse - before <- mean((data_beforeFertility - data_beforepred)^2)$

$mse - after <- mean((data_afterFertility - data_afterpred)^2)$