Spatial analysis of maternal mortality patterns and associated factors in Africa

Tsoanelo Irene Masilo (tsoanelo@aims.ac.za)
African Institute for Mathematical Sciences (AIMS)

Supervised by: Dr. Siaka Lougue
Institut de recherche en sciences de la santé, Burkina Faso

23 May 2019

Submitted in partial fulfillment of a structured masters degree at AIMS South Africa

Abstract

There are many strategies implemented to reduce the maternal mortality ratio in Africa but it still appears to be a public health challenge. In this study we observe the ratios of maternal mortality in African countries over the period 2011-2015. The aim is to observe if maternal mortality disparities exist between African countries, and if they do exist identify potential causes. We took into account health and socio-economic factors known to have an effect on maternal mortality. The data was first tested for normality using the Shapiro-Wilk test. Spatial analysis methods were used to map maternal mortality by countries to see if ratios of countries are similar or dispersed. Kendall's Tau correlation test was used to determine the relationship of each indicator with maternal mortality. Rather than using a typical panel data regression model as in other studies, this study considered the type of the dependent variable. A panel data fractional response regression model was used to asses the relationship between associated factors and maternal mortality in Africa.

Declaration

I, the undersigned, hereby declare that the work contained in this research project is my original work, and that any work done by others or by myself previously has been acknowledged and referenced accordingly.

Tsoanelo Irene Masilo, 23 May 2019

Contents

1. Introduction

1.1 Background

Maternal mortality has always been and continues to be a public health crisis all over the world and unacceptably high in Africa. Even though Africa has high maternal mortality, there are large disparities within the continent; there are some countries that have low ratios of maternal mortality and others contributing a lot to Africa's maternal mortality ratio. In the International Statistical Classification of Diseases and Related Health Problems the World health organisation (WHO, 2007) defines maternal mortality as the death of a pregnant woman, during child birth or within 42 days of termination of pregnancy, from any cause related to or aggravated by the pregnancy not from accidental or incidental causes. Maternal mortality ratio is calculated as the number of maternal deaths in a population divided by the number of live births in the same year multiplied by 100 000 (WHO, 2007).

The WHO fact sheet updated in 2018 stated that about 830 women die every day around the world due to pregnancy related complications or during childbirth and 99% of these deaths occur in developing countries. In the year 2015, it was estimated that the maternal mortality ratio in developing countries was 239 deaths per 100 000 live births while in developed countries it was 12 per 100 000 (WHO, 2018). The lifetime risk of maternal death was estimated as 1 in 3 300 in high income countries while in low income countries it was estimated as 1 in 41 (UNICEF, 2017). Maternal mortality ratios are high in poor countries than in rich countries. One of the main issues is health care for pregnant women. WHO recommends a minimum of four antenatal care visits but only about half of all pregnant women receive the recommended amount of care. Women in Sub-Saharan Africa and South Asia that attend at least four antenatal visits are 52% and 46%, respectively (UNICEF, 2018). It was then discovered that in regions with high maternal mortality ratio, few women had received the recommended amount of care.

There are many organisations working to reduce maternal mortality globally. Maternal death surveillance and response (MDSR), provides information to help countries to review maternal deaths enabling them to plan interventions to prevent these deaths in future (WHO, 2013). One target of the Sustainable Development Goals is to decrease the global maternal mortality ratio to less than 70 per 100 000 births, no country should have two times more than the global average maternal mortality ratio by 2030. The fifth millennium development goal of the United Nations is to better maternal health, its target was to reduce the maternal mortality ratio by 75% between the years 1990 and 2015 (WHO, 2015). From 1990-2015 the global maternal mortality ratio decreased by 44%, that is; from 385 to 216 deaths per 100 000 births. Even though it is a notable decrease, still the target was not reached (UNICEF, 2017). Most of maternal deaths are said to be preventable but maternal mortality still remains a challenge, in Africa.

1.2 Statement of the problem

Johnson, highlights that the rate at which women die during childbirth is alarming. It is very sad how maternal mortality leaves children motherless and husbands as widows even worse the newborn child might not survive sometimes. In one study, it was discovered that children whom their mothers died are more likely to die than those who lost fathers and those who still had both parents (National Academies Press, 2000). Maternal mortality highly contributes to an increased number of orphans in the community and most homes become single-parent households.

1.3 Objectives

Main objective: To study patterns of maternal mortality in Africa and to explore associated factors in order to help reduce maternal deaths.

Maternal mortality ratios between and within African countries are compared over time to identify if disparities exist. Identify what countries with high maternal mortality ratios can learn from countries with low maternal mortality. The objective can be achieved by understanding associated factors and observing how they affect countries of different levels of maternal mortality.

1.4 Research questions

Maternal mortality ratios in Africa differ from one country to another. This study desires to answer the following questions;

1. Are there significant disparities of maternal mortality among African countries?
2. What can explain these disparities among African countries?
3. Which countries have high ratios and which have low ratios of maternal mortality?
4. Which factors are significantly associated with maternal mortality?

1.5 Organisation of the essay

This essay is categorised into five chapters to accomplish the aim of the study. **Chapter 1** introduces maternal mortality as a problem in African countries and provides some background on maternal mortality. It includes the problem statement, the objective of conducting the study and research questions it aims to answer. **The second chapter 2** entails the literature review it covers findings and views of other researchers who studied maternal mortality before. **Chapter 3** is the methodology part, it describes the data used and how it was extracted. It provides discussions on methods of analysis appropriate for the study objectives. **The fourth chapter 4** contains results and discussions obtained in the analysis. Lastly, conclude and give recommendations based on what the results portray in **chapter 5**.

2. Literature Review

In this chapter, past studies by other researchers on maternal mortality are reviewed. The reviews are based on different factors associated with maternal mortality with particular focus on African countries.

According to WHO, the leading causes of maternal deaths are; high blood pressure during pregnancy (pre-eclampsia and eclampsia), severe bleeding usually after childbirth, infections, unsafe pregnancy termination and complications due to delivery. Maternal mortality or morbidity does not depend only on those, Murray (2019) points out that an area where one lives can also be a contributing factor to maternal mortality, it has a noticeable effect on the health and well-being of a pregnant woman living there.

Rana et al. (2018), conducted a study that provides how the relationship between health expenditure and health care outcomes may differ across countries because of different income levels.They reported that; globally in the past few years health expenditure has been rising, and that since 1990 to 2013, the reduction in maternal mortality ratio was 45% leading to an increase of life expectancy at birth from 64 years to 71 years. They concluded that increasing health expenditure has a potential to improve maternal health in women.

To decrease the ratio of maternal mortality women need to have access to health care throughout pregnancy, during the time of delivery and after. Better access to quality health care services and presence of skilled physicians can improve maternal health, Johnson. Antenatal care reduces complications during pregnancy and thus reduces the risk of maternal mortality. Candace Johnson reported that highest MMRs are found in low income countries and lowest ratios in high income countries. Women in high income countries attend at least four antenatal care visits and receive care during childbirth and post partum, while in low income countries only a few attend antenatal care for at least four visits

It is within people's knowledge that teenage pregnancy carries high risk of maternal mortality, Girum and Wasie (2017), early pregnancy may lead to early child bearing and because of the less developed body this could cause maternal mortality in teenage girls. At young and old ages the risk of maternal mortality is high.The risk is lesser for middle ages. A study to find out whether there is a relationship between number of children a woman has and maternal mortality was conducted, Koski-Rahikkala et al. (2006). They found that mortality was lowest among women who have two to four children, this means having a high number of children can lead to an increased risk of maternal mortality.

In one study it was reported that differentials in maternal mortality by socio-economic status do exist among and within countries. They also concluded that people from poor and disadvantaged countries are more likely to die than those that are not poor, McCarthy and Maine (1992). In a research article by Ensor et al. (2010), in the analysis table of associations between health outcomes and changes in GDP per capita, they reported that a decrease in GDP per capita led to an increase of the maternal mortality ratio.

A study was conducted to evaluate the changes in human development index, maternal mortality rate and under five mortality rate from 1980 to 2010 in some West Asian countries and also to study the relationship between these indicators, Alimohamadi et al. (2017). It showed that Humam development index (HDI) has a negative relationship with Maternal mortality thus an increase in HDI leads to a decrease in maternal mortality. On the other hand, high levels of education have also been associated with better maternal health, lack of education is a barrier for women to access healthcare Diorio and Dr.Crivelli-Kovach (2014). New UNESCO's analysis proves that education transforms development, it was reported that education can reduce maternal deaths by helping women to recognize the importance

of health care services, they should seek care and make sure that during delivery they are attended by trained health workers.

A study was conducted by Berhan and Berhan (2014) to assess the relationship of perinatal mortality rates and maternal mortality ratios with the proportion of skilled health personnel who assisted in deliveries. In the study it was found that reduction of maternal mortality was significantly associated with an increase in skilled delivery attendance. They recommended that countries with high MMR need to benchmark the experience of those with a lower MMR to reduce the high maternal deaths.

Delivery by caesarean section is an effective way of birth delivery to save maternal lives, but once the rate of caesarean section goes higher than 10% at population level, it does not necessarily decrease maternal mortality rates. Thus high rate of caesarian sections may not be a good intervention for achieving the lowest maternal mortality, Ye et al. (2015).

Alvarez et al. (2009), conducted one study to describe and determine different factors associated with maternal mortality in Sub-Saharan countries using data collected between 1997 and 2006. They discovered that there was a negative relationship between maternal mortality ratio and antenatal care coverage, births assisted by skilled health worker, adult literacy, gross national income per capita and expenditure on health. Meh (2017) in one study pointed out that factors like age, parity and access to health care affect maternal mortality. Bayati et al. (2016), conducted a study to identify the effect of associated factors on maternal mortality in Eastern Mediterranean region. They did a panel data analysis of 2004-2011 and applied the fixed effects model to estimate the parameters. Their results also showed that there is a significant negative relationship between maternal mortality and GDP per capita, female literacy rate, health expenditure and skilled birth attendance. Girum and Wasie (2017) recommended that policies and programs targeted to reduce maternal mortality should look into things like population dynamics, socio-economic and health system factors that have an influence on maternal health.

3. Methodology

3.1 The data

In this essay secondary data was used, it was obtained freely from the internet; from the World Bank database[1], Our World in Data database[2] and the United Nations database[3]. We also extracted some data from individual country profiles for those countries that data was unavailable. We extracted information for the years 2011-2015 on the 54 African countries. A panel data approach is used, whereby the 54 countries are studied over a period of five years (2011-2015).

To deal with incomplete data we used methods of data interpolation to fill in the missing values. We estimated the missing data values by taking the average of values in the neighbourhood. If a data point of a certain year is unavailable we averaged data values of the year before and the year after that certain year, that is; $(b+a)/2$, a and b respectively representing data value of the year before and the year after . In this study we specifically considered data of the five year period 2011-2015 due to absence of the latest data. We also considered that at country level there is no significant change within a five years time interval, so if a country had no information for the the other years we imputed no change. Where we did not have values for the study period we used the most recent available data value to impute for the other years.

We used the modelled estimates of maternal mortality ratio (MMR) per 100,000 live births as the dependent (response) variable. We also used some of the indicators known to be associated with MMR as independent (predictor) variables. Predictor variables used are the following; current health expenditure (CHE), gross domestic product per capita (GDPpc), human development index (HDI), total fertility rate (TFR), antenatal care coverage of at least four visits (ANC4), health access and quality index (HAQ), physicians per 1000 people (MED.PHYS),female literacy rate (LR.F) and births attended by skilled health staff (BA). Definitions of these variables in the table below are obtained from the mentioned data sources.

Table 3.1: **Definitions of the variables used in the study**

Indicator	Definition
MMR	The number of women who die due to pregnancy-related causes while pregnant or within forty two days of pregnancy termination per 100,000 live births.
CHE	It is expressed as a percentage of GDP. The estimates includes healthcare goods and services used throughout a year. It does not include capital health expenditures such as buildings, machinery, IT and stocks of vaccines for emergency or outbreaks.
GDPpc	It is the GDP divided by midyear population and it is given in U.S dollars. GDP is the sum of gross value added by all producers in the economy plus product taxes minus any subsidies not included in the value of the products. When calculating it no deductions made for depreciation of fabricated assets or for depletion and degradation of natural resources.
HDI	A composite index that measures the average achievement in three basic dimensions of human development; a long and healthy life, knowledge and a decent standard of living.

[1]Source: https://data.worldbank.org[Last Accessed March 2019]
[2]Source: https://ourworldindata.org [Last Accessed March 2019]
[3]Source: https://data.un.org [Last Accessed March 2019]

TFR	The number of children that would be born to a woman if she were to live to the end of her childbearing years giving birth to children in accordance with age-specific fertility rates of the specified year.
ANC4	Percentage of women who attended healthcare services at least four times throughout pregnancy by skilled health personnel for reasons related to pregnancy.
HAQ	An index measured on a scale from 0 (worst) to 100 (best) based on death rates from causes that could have be avoided by timely and effective medical care
MED.PHYS	Includes all generalist and specialist medical practitioners in the country.
LR.F	It is the percentage of women aged 15 and above who can read, write and understand short simple statements in their daily life.
BA	Percentage of birth deliveries attended by personnels trained to provide the necessary supervision, care and advice to women throughout pregnancy, labour and the post-partum period; trained to conduct deliveries on their own and care for newborns.

3.2 Descriptive statistics

Before going deep into analysis the data should be summarised to easily understand it. The summary statistics table is used to show quartiles, minimum, maximum and mean values of variables. To know how to proceed with the analysis the distribution of the data has to be identified using tests of normality.

3.2.1 Shapiro-Wilk test. The Shapiro-Wilk test is a tool used to test if a population comes from a normal distribution. It tests the hypothesis that a population belongs to a normal distribution.

H_0 : The data are normally distributed.
H_a : The data are not normally distributed.

The Shapiro-Wilk test statistic is given by:

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}, \tag{3.2.1}$$

where:
n- the number of observations,
x_i- the i^{th} data value of the original data,
$\bar{x}$- the mean of the original data,
$x(i)$- the i^{th} data value of the ordered data,
a_i- the i^{th} component of the vector $(a_1, a_2, \ldots, a_n)$,

$$(a_1, a_2, \ldots, a_n) = \frac{m'V^{-1}}{(m'V^{-1}V^{-1}m)^{1/2}}. \tag{3.2.2}$$

where:

$$m = \begin{pmatrix} m_1 \\ m_2 \\ . \\ . \\ . \\ m_n \end{pmatrix}, \quad m_i = E(x_{(i)}),$$

$$V = \begin{pmatrix} v_{11} & \dots & v_{1n} \\ \vdots & \ddots & \\ v_{n1} & & v_{nn} \end{pmatrix}, \quad v_{ij} = cov(x_{(i)}, x_{(j)}).$$

Testing at 5% level of significance ($\alpha = 0.05$), the null hypothesis is rejected if the statistic W is less than the critical value W_α and the p-value is less than α. Concluding that at 95% confidence the data is not normally distributed. Otherwise fail to reject the null hypothesis and conclude with 95% confidence that the data is normally distributed.

3.2.2 Kendall's Tau correlation coefficient. Kendall's Tau correlation coefficient is a tool that measures the strength of the association between two random variables. Its values are in the range [-1,1]. A negative correlation coefficient indicates that when the rank of the other variable increases the rank of the other decreases. A positive correlation coefficient indicates that the ranks of both the variables move to the same direction, if the rank of the other variable increases and so does the rank of the other variable. If the coefficient is zero it means the variables are not associated.

Kendall's Tau is calculated from ranks of the points (X, Y). Let U represent the variable of ranks of X and V the ranks of Y.

The points (U_i, V_i) and (U_j, V_j) are said to be concordant if

$$(U_i < U_j) \text{ and } (V_i < V_j)$$

OR

$$(U_i > U_j) \text{ and } (V_i > V_j).$$

The points are said to be discordant if

$$(U_i < U_j) \text{ and } (V_i > V_j)$$

OR

$$(U_i > U_j) \text{ and } (V_i < V_j).$$

Kendall's Tau is given by:

$$\tau = \frac{2(n_c - n_d)}{\sqrt{N(N-1) - T_X}\sqrt{N(N-1) - T_Y}}. \tag{3.2.3}$$

$$T_X = \sum_{i=1}^{S_X} (t_{(X)i}^2 - t_{(X)i}). \tag{3.2.4}$$

$$T_Y = \sum_{i=1}^{S_Y} (t_{(Y)i}^2 - t_{(Y)i}). \tag{3.2.5}$$

Where:
N- number of observations,
n_c - total number of concordant pairs,
n_d - total number of discordant pairs,
S_X- number of sets of ties in the X variable,
S_Y- number of sets of ties in the Y variable,
$t_{(X)i}$- number of ties in the i^{th} set of ties of X,
$t_{(Y)i}$- number of ties in the i^{th} set of ties of Y.

3.3 Disease mapping

Disease mapping is used to describe the spatial patterns in disease risk across different areas to identify those that have high risk of the disease, Lee et al. (2014). We applied techniques of disease mapping using R software to visually represent the data on the main problem; maternal mortality in African countries.

After visualising the spatial patterns we then had to analyse using spatial analysis. Spatial analysis is applied mostly in the analysis of geographic data so we made use of it to describe maternal mortality ratios of countries and their variations from one country to the other.

The spatial phenomenon for the area data is said to be represented as:

$$\{Y(A_i), A_i \in A_1, A_2, \ldots, A_n\},$$

where: $(A_1, A_2, \ldots, A_n)$ are sub-groups of the study region R, $R = A_1 \cup A_2 \cup \cdots \cup A_n$.

3.4 Spatial autocorrelation

Spatial autocorrelation is used to measure the relationship of values of the same variable at various locations. Some of the test statistics used to test for spatial autocorrelation are Moran's I and Geary's C. Below is the brief discussion of these measures of spatial autocorrelation, more details discussed on Moran's I which is used in the essay.

3.4.1 Moran's I index. Moran's I is a statistical measure of spatial autocorrelation which was developed by Patrick Alfred Pierce Moran.It evaluates if the spatial patterns of the disease or mortality are clustered, dispersed or random. The Moran's I index is given by:

$$I = \frac{N}{S_0} \frac{\sum_i^N \sum_j^N w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i^N (x_i - \bar{x})^2}, \; j \neq i, \tag{3.4.1}$$

where:
N- number of spatial units indexed i and j,
x- the variable of interest,
$\bar{x}$- the mean of x,
x_i, x_j-variable values of a spatial units i and j respectively,
w_{ij}- matrix of spatial weight where $w_{ii} = 0$,
S_0- sum of w_{ij}‘s.

The expected value and variance of the Moran's I is given by:

$$E(I) = \frac{-1}{N-1}, \tag{3.4.2}$$

and

$$Var(I) = \frac{NS_4 - S_3S_5}{(N-1)(N-2)(N-3)S_0^2} - E(I)^2, \tag{3.4.3}$$

where;

$$S_1 = \frac{1}{2}\sum_i^N\sum_j^N(w_{ij} + w_{ji})^2, \tag{3.4.4}$$

$$S_2 = \sum_i^N \left(\sum_j^N w_{ij} + \sum_j^N w_{ji}\right)^2, \tag{3.4.5}$$

$$S_3 = \frac{N^{-1}\sum_i^N(x_i - \bar{x})^4}{(N^{-1}\sum_i^N(x_i - \bar{x})^2)^2}, \tag{3.4.6}$$

$$S_4 = (N^2 - 3N + 3)S_1 - NS_2 + 3S_0, \tag{3.4.7}$$

$$S_5 = (N^2 - N)S_1 - 2NS_2 + 6S_0. \tag{3.4.8}$$

Values of Moran's I index fall in the range [-1,1]. Positive values indicate a positive autocorrelation, that is the values of the neighbouring locations tend to be the same, meaning that the pattern is spatially clustered. Negative values indicate a negative autocorrelation, the values of the compared neighbouring locations tend to be different, meaning the pattern is spatially dispersed. If the value of I is closer to zero it indicates no significant spatial autocorrelation, values are spatially random. Moran's I is an inferential statistics so its values can not be directly interpreted but can be interpreted with the use of hypothesis testing.

Testing the hypothesis;

H_0 : values are spatially independent.
H_a : values are not spatially indepedent.

The Moran's I standardised value is given by:

$$Z_I = \frac{I - E(I)}{\sqrt{Var(I)}}, \tag{3.4.9}$$

with the associated p-value at 5% level of significance($\alpha = 0.05$) or at 10% level of significance ($\alpha = 0.1$). If the value of the Z score is greater than the critical value the null hypothesis is rejected at 5% (or 10% if $\alpha = 0.1$) level of confidence and conclude there is a spatial autocorrelation. Otherwise if the value of the Z score is less than the critical value fail to reject the null hypothesis and conclude that there is no spatial autocorrelation.

Moran's scatter plot. We can also visualise the spatial autocorrelation using the Moran's scatter plot, it plots the spatial data against its lagged values (average value for the neighbourhood spatial data). The x-axis units of the plot can be in the form of deviations from the mean or the original values of the variable. The y-axis can be in the form of standardized weights or values of the lagged variable. The slope of the scatter plot is the moran's I statistic.

The Moran scatter plot has four quadrants representing high-high (upper-right), low-low (lower-left), high-low (lower-right) and low-high (upper-left). High-high and low-low are entities with a positive autocorrelation while high-low and low-high are those negatively autocorrelated. In the high-high quadrant we have entities with high ratios surrounded by entities with high ratios of the variable of interest. In the low-low quadrant we have entities with low ratios surrounded by entities with low ratios. In the high-low and low-high quadrants there is no correlation other entities have high ratios others low ratios.

3.4.2 Geary's C index. Geary's C also known as Geary's contiguity ratio was developed by Roy C. Geary. It is a measure of spatial autocorrelation that is inversely related to Moran's I. Unlike Moran's I, Geary's C is more sensitive to local spatial autocorrelation.

Geary's C is given by:

$$C = \frac{(N-1)\sum_{i}^{N}\sum_{j}^{N} w_{ij}(x_i - x_j)^2}{2S_0 \sum_{i}^{N}(x_i - \bar{x})^2}, \; j \neq i. \tag{3.4.10}$$

where variables are defined as in the case of Moran's I. The expected value of Geary's C is 1. A value of C greater than 1 indicates a negative spatial autocorrelation while a value less than 1 indicates a positive spatial autocorrelation.

3.5 Panel data regression

Panel regression is one of the methods used in modelling panel data. This is the kind of data that has repeated observational units over time, it is also known as cross-sectional time-series data or longitudinal data. In panel data there can be certain aspects that change over time, errors can be correlated for a particular entity and the effects on the entities may be different (existence of heterogeneity). Panel regression accounts for these issues when analysing. Dealing with these kind of data we can not apply the pooled ordinary least squares (OLS) that is commonly used in regression because it neglects its time-series and cross sectional nature, Hossain (2013). OLS assumes that effects on entities are similar and that the errors are not correlated, hence why we introduced panel data regression in this study as the efficient method to use.

In this study there are multiple predictor variables and a numerical response variable, so we focus more on the multivariate approach of panel data. In multivariate regression the model simultaneously estimates more than one regression equation.

The general panel regression model is given by:

$$y_{it} = \beta_0 + \beta_1 x_{1,it} + \ldots + \beta_k x_{k,it} + u_{it}, \tag{3.5.1}$$

which can be represented as:

$$y_{it} = \beta_0 + \boldsymbol{\beta}\mathbf{X}_{it} + u_{it}, \tag{3.5.2}$$

where:
i- entity $(i = 1, ..., N)$,
t - time $(t = 1, ..., T)$,
y_{it} - the dependent variable for entity i at time t,
β_0 - the intercept,
$\boldsymbol{\beta}$ - a column vector of coefficients $\beta_1, \ldots, \beta_k$,
$\mathbf{X}$ - a row vector of the predictor variables $x_1, \ldots, x_k$,
u_{it} - the error term.

In panel data analysis we have two main kinds of regression models that are used to explore the relationship between the response variable and a set of predictor variables within an entity (entities are countries in our case). The models are; fixed effects model and random effects model.

3.5.1 Fixed effects model. Fixed effects is used to investigate the impact of variables that change over time, it does not estimate time-invariant effects. Using this model we make an assumption that variations across entities are correlated to the predictor variables. The model does not estimate is given as:

$$y_{it} = \beta_0 + \boldsymbol{\beta}\mathbf{X}_{it} + \delta Z_i + u_{it}, \tag{3.5.3}$$

$\alpha_i = \delta Z_i + u_{it}$ then:

$$y_{it} = \alpha_i + \boldsymbol{\beta}\mathbf{X}_{it} + u_{it}, \tag{3.5.4}$$

where:
Z_i-the unobserved time invariant heterogeneities across entities.
α_i - intercept that is in terms of the unobserved time invariant effect.
δ-a coefficient.

Notice that α_i is not observed so the fixed effects estimator eliminates it by fully demeaning the variables by the 'within transformation'. Then have:

$$y_{it} - \bar{y_i} = \boldsymbol{\beta}(\mathbf{X}_{it} - \bar{\mathbf{X}}_i) + (\alpha_i - \bar{\alpha_i}) + (u_{it} - \bar{u_i}), \tag{3.5.5}$$

where:

$$\bar{y_i} = \frac{1}{N}\sum_{i=1}^{N} y_{it} \ , \quad \bar{\mathbf{X}}_i = \boldsymbol{\beta}\frac{1}{N}\sum_{i=1}^{N}\mathbf{X}_{it} \ , \quad \bar{u_i} = \frac{1}{N}\sum_{i=1}^{N} u_{it} \ , \quad \bar{\alpha_i} = \frac{1}{N}\sum_{i=1}^{N}\alpha_i.$$

Because α_i is a constant, $\alpha_i = \bar{\alpha_i}$ hence the resulting estimation where the unobserved effect is eliminated is:

$$y_{it} - \bar{y_i} = \boldsymbol{\beta}(\mathbf{X}_{it} - \bar{\mathbf{X}}_i) + (u_{it} - \bar{u_i}). \tag{3.5.6}$$

The pooled Ordinary Least Square is applied to the demeaned data to provide the fixed-effects estimator β. An alternative way to run a fixed effects regression model can be applying the pooled Ordinary Least Square to a panel model with added $N-1$ dummy variables;

$$y_{it} = \beta_0 + \boldsymbol{\beta}\mathbf{X}_{it} + \gamma_2 D2_i + \ldots + \gamma_N DN_i + u_{it}. \tag{3.5.7}$$

3.5.2 Random effects model. This model as opposed to fixed effects model variation across entities is assumed not to be correlated with the predictor variables. The model estimates the effects of time-invariant variables. It is given by

$$y_{it} = \alpha + \boldsymbol{\beta}\mathbf{X}_{it} + \alpha_i + u_{it}. \tag{3.5.8}$$

The random effects estimator partially demeans the variables to have the quasi-demeaned data:

$$y_{it} - \theta\bar{y_i} = \alpha(1-\theta) + \boldsymbol{\beta}(\mathbf{X}_{it} - \theta\bar{\mathbf{X}}_i) + (\alpha_i - \theta\bar{\alpha_i}) + (u_{it} - \theta\bar{u_i}), \tag{3.5.9}$$

where:

$$\theta = 1 - \sqrt{\frac{\sigma_u^2}{\sigma_u^2 + T\sigma_\alpha^2}}. \tag{3.5.10}$$

If θ=0 then it gets closer to the pooled ordinary least squares estimator, that is when the variance of the unobserved effect σ_α^2 gets smaller. If θ =1 then we have the fixed effects estimator. When T and/or σ_α^2 get large the random effects estimator gets closer to the fixed effect estimator.

3.5.3 Durbin-Wu-Hausman test. Between the fixed effects regression estimator and the random effects regression estimator which one is the best for our data? To know which estimator is efficient between fixed effects and random effects the Hausman test[4] is used. It is used to check if the errors are correlated with the predictor variables.

If there is no correlation: $cov(\alpha_i, x_{it}) = 0$
$\implies$ β_{RE} and β_{FE} are consistent.
$\implies$ $SE(\beta_{RE}) < SE(\beta_{FE})$, β_{RE} is more efficient.
We then conclude in using random effects estimator.

If there is correlation: $cov(\alpha_i, x_{it}) \neq 0$
$\implies$ β_{FE} is consistent and efficient.
We then conclude in using the fixed effects estimator.

β_{FE} is the fixed effects model coefficient and β_{RE} is the random effects model coefficient.

We test the hypothesis:

H_0 : use random effects estimator.
H_a : use fixed effects estimator.

The Hausman test statistic is given by:

$$H = \frac{(\beta_{FE} - \beta_{RE})^2}{var(\beta_{FE}) - var(\beta_{RE})}. \tag{3.5.11}$$

H is distributed with χ_k^2 where k is the number of predictors. To carry on with the test we first estimate the fixed effects and random effects estimators and use them to run the Hausman test. If the value of

[4]Source: http://en.m.wikipedia.org/wiki/Durbin-Wu-Hausman_test (accessed on 18^{th} April2019)

H is larger than the critical value and the p-value is less than the level of significance; we reject the null hypothesis and use the fixed effects. Otherwise if H is smaller and the p-value is large we use random effects.

3.6 Fractional regression estimation

The topic of fractional regression (FREG) models was first suggested by Papke and Wooldridge (1996), in one of their studies; "voluntary individual contributions to retirement accounts". FREG models are used when all possible values of the response variable are fraction types of information meaning they fall in the interval $[0, 1]$, hence why we refer to those variables as fractional response variables. When y is bounded by zero and one we can not be certain to use a linear model (Calabrese, 2012) thus we will use a non linear models in fractional regression, a model of the form:

$$\hat{y} = f(\alpha + \beta x). \tag{3.6.1}$$

When the variable represents values like ratios, proportions or percentages we then use fractional response regression that estimates the conditional expected value of the response variable. We use fractional regression to examine how the predictor variables influence the response variable, capturing the non linearity nature of the data. We fit a regression for the mean of y (response) conditional on $x's$ (predictors).

The standard FREG model is given by:

$$E(y_i \mid \mathbf{X_i}) = G(\boldsymbol{X_i\beta}). \tag{3.6.2}$$

Where:

$$G(\boldsymbol{X_i\beta}) = \frac{exp(\boldsymbol{X_i\beta})}{1 + exp(\boldsymbol{X_i\beta})}, \qquad \text{for a fractional logit model.} \tag{3.6.3}$$

$$G(\boldsymbol{X_i\beta}) = \Phi(\boldsymbol{X_i\beta}), \qquad \text{for a fractional probit model.} \tag{3.6.4}$$

$$G(\boldsymbol{X_i\beta}) = \Phi\left(\frac{\boldsymbol{X_i\beta}}{exp(\boldsymbol{Z_i\gamma})}\right), \qquad \text{for a fractional heteroskedastic probit model.} \tag{3.6.5}$$

and:
$0 < G(\boldsymbol{X_i\beta}) < 1$ is the link function.
Φ : standard normal cumulative density function.
$\boldsymbol{X_i}$: a vector of predictor variables.
$\boldsymbol{Z_i}$: covariates to model variance.
y_i : response variable of the i^{th} observation. $0 \leq y_i \leq 1$.
$\boldsymbol{\beta}$,$\boldsymbol{\gamma}$: vectors of parameters of interest.

Liu and Xin (2014) explain that we can not consider a fractional response variable to be normally distributed because its values are not defined on the whole real line. Due to this, in fractional response regression the Quasi maximum likelihood estimation (QMLE) methods are used. The authors continue to add that in QMLE there is no need to assume any distribution which is why is the best approach.

QMLE allows for the direct estimation of values at the extremes $(y = 0 \; and \; y = 1)$ without an Ad-hoc transformation.

So to estimate the parameters; the Bernoulli log-likelihood function below:

$$l_i(\boldsymbol{\beta}) \equiv y_i \log[G(\mathbf{X_i}\beta)] + (1 - y_i) \log[1 - G(\mathbf{X_i}\beta)], \tag{3.6.6}$$

is maximized, the Quasi maximum likelihood estimator (QMLE) for $\boldsymbol{\beta}$

$$\hat{\boldsymbol{\beta}} = \max \sum_{1}^{N} l_i(\boldsymbol{\beta}).$$

If $E(y_i \mid \mathbf{X_i})$ is specified correctly the QMLE is said to be consistent and asymptotically normal, regardless of the distribution of y conditional on the $x's$. Thus the response variable can have continuous and/or discrete characteristics Gallani et al. (2015).

Hausman and Leonard (1997) developed a fixed effects logit QMLE for panel data, specifically for long panel data. Ramalho et al. (2016) points out that the model has a weakness of not allowing for time invariant effects. Papke and Wooldridge (2008) also proposed an extension for panel data to allow for unobserved, time-invariant effects. They came up with correlated random effects probit estimators. They used the probit formulation of the fractional regression model given as:

$$E(y_{it} \mid \alpha_i, \mathbf{X_{it}}) = G(\boldsymbol{X_{it}}\boldsymbol{\beta} + \alpha_i), \tag{3.6.7}$$

where variables are defined as before but now with an added variable for the time period t and α_i unobserved time invariant effect.

4. Results and Discussions

In the analysis of this study we used R software to run methods explained in chapter three.

4.1 Descriptive statistics

Table 4.1: **Shapiro-Wilk test results of normality of maternal mortality for the years 2011-2015**

Shapiro-Wilk test						
Variable	statistic	2011	2012	2013	2014	2015
Maternal mortality ratio	W	0.9431	0.9471	0.9494	0.9519	0.9553
	p-value	0.0126	0.0187	0.0234	0.0301	0.0427
Current health expenditure	W	0.9503	0.9664	0.9602	0.8094	0.8294
	p-value	0.0256	0.1335	0.0701	6.869e-07	2.208e-06
Gross domestic product	W	0.6323	0.6372	0.6522	0.6572	0.6837
	p-value	2.238e-10	2.683e-10	4.761e-10	5.784e-10	1.67e-09
Human development index	W	0.9555	0.9555	0.9607	0.9645	0.9672
	p-value	0.0433	0.0433	0.0743	0.1093	0.1446
Health access & quality index	W	0.9133	0.9133	0.9209	0.9290	0.9290
	p-value	0.0008	0.0008	0.0016	0.0033	0.0033
Antenatal care (4visit)	W	0.9717	0.97154	0.9697	0.9705	0.9686
	p-value	0.2595	0.2560	0.2156	0.2326	0.1934
Total fertility rate	W	0.9701	0.9727	0.9751	0.9768	0.9786
	p-value	0.1939	0.2515	0.3186	0.3756	0.4410
Female literacy rate	W	0.9482	0.9440	0.9392	0.9401	0.9375
	p-value	0.0207	0.0137	0.0086	0.0094	0.0073
Physician density	W	0.6016	0.6480	0.6570	0.6476	0.6440
	p-value	9.466e-11	5.145e-10	7.265e-10	5.079e-10	4.423e-10
Births by skilled	W	0.9664	0.9630	0.9623	0.9504	0.9508
	p-value	0.1328	0.0935	0.0875	0.0258	0.0270

For all the five years the p-values for maternal mortality ratio, gross domestic product, health access and quality index, female literacy rate and physician density are significant (less than $\alpha = 0.05$). The significant p-values indicate that the data of these variables come from a non-normal distribution. Total fertility rate and percentage of women who visited antenatal care at least four times have insignificant p-values (greater than $\alpha = 0.05$) throughout all the years thus they both follow a normal distribution.

Current health expenditure data followed a normal distribution for the years 2011,2014 and 2015 and for the years 2012 and 2013 the distribution was normal. Human development index show normality for the years 2011 and 2012 and then non normality for the years 2013,2014 and 2015. Births attended by skilled personnel data is normally distributed for the first three years and not normally distributed in 2014 and 2015.

Table 4.2: **The summary statistics of variables in the year 2011**

Variables	Minimum	1st quartile	Median	Mean	3rd quartile	Maximum
Maternal mortality ratio	9	315.5	428.0	477.9	647.8	1580.0
Current health expenditure	1.349	3.924	5.138	5.564	6.447	10.555
Gross domestic product	247.2	584.4	1324.3	2706.7	3157.5	21451.9
Human development index	0.2850	0.4285	0.4995	0.5179	0.5723	0.7700
Health access and quality index	28.80	40.33	43.05	45.39	48.55	69.70
Antenatal care(4visits)	6.30	44.45	55.50	55.31	72.65	87.10
Total fertility rate	1.550	3.666	4.928	4.604	5.315	7.455
Female Literacy rate	12.19	32.99	59.20	56.03	76.81	94.45
Physician density	0.0080	0.0500	0.1000	0.3246	0.3740	2.8300
Births by skilled	10.00	50.25	65.50	65.47	82.15	99.80

Table 4.3: **The summary statistics of variables in the year 2012**

Variables	Minimum	1st quartile	Median	Mean	3rd quartile	Maximum
Maternal mortality ratio	9	302.8	415.5	465.4	630.0	1510.0
Current health expenditure	1.445	3.957	4.989	5.431	7.089	9.283
Gross domestic product	250.4	612.1	1266.2	2829.7	3124.1	21557.7
Human development index	0.2850	0.4285	0.4995	0.5179	0.5723	0.7700
Health access and quality index	28.80	40.33	43.05	45.39	48.55	69.70
Antenatal care(4visits)	6.30	44.65	55.50	55.59	73.85	87.10
Total fertility rate	1.540	3.643	4.872	4.547	5.245	7.420
Female Literacy rate	12.19	32.99	60.55	56.38	76.81	94.45
Physician density	0.019	0.053	0.099	0.309	0.384	1.996
Births by skilled	16.55	54.75	65.50	65.65	82.15	99.80

Table 4.4: **The summary statistics of variables in the year 2013**

Variables	Minimum	1st quartile	Median	Mean	3rd quartile	Maximum
Maternal mortality ratio	9	289.2	411.5	453.5	631.8	1460.0
Current health expenditure	1.725	4.046	5.094	5.557	6.794	11.579
Gross domestic product	255.4	691.6	1339.7	2858.6	3188.0	20247.0
Human development index	0.2850	0.4348	0.5060	0.5231	0.5817	0.7780
Health access and quality index	28.70	40.50	44.12	46.38	49.73	69.80
Antenatal care(4visits)	6.30	44.55	55.70	56.08	73.85	87.10
Total fertility rate	1.440	3.612	4.793	4.485	5.180	7.381
Physician density	0.0190	0.0530	0.0960	0.3062	0.3840	2.0440
Female mean age of bearing	27.43	28.50	29.04	29.27	29.84	32.85
Births by skilled	18.40	54.08	67.75	66.87	84.41	99.90

Table 4.5: **The summary statistics of variables in the year 2014**

Variables	Minimum	1st quartile	Median	Mean	3rd quartile	Maximum
Maternal mortality ratio	9	278.5	400.5	442.6	624.0	1410.0
Current health expenditure	1.775	4.154	5.338	5.830	6.811	19.727
Gross domestic product	273.5	703.1	1330.9	2793.2	3301.2	19245.7
Human development index	0.2850	0.4412	0.5090	0.5281	0.5875	0.7860
Health access and quality index	28.60	41.67	45.05	47.37	50.85	70.10
Antenatal care(4visits)	6.30	45.90	56.60	56.82	73.85	87.30
Total fertility rate	1.430	3.572	4.697	4.422	5.114	7.338
Female Literacy rate	13.93	33.58	60.55	57.02	76.81	94.45
Physician density	0.0190	0.0530	0.0950	0.3077	0.3950	2.0920
Births by skilled	15.50	55.02	72.20	68.31	88.00	99.80

Table 4.6: **The summary statistics of variables in the year 2015**

Variables	Minimum	1st quartile	Median	Mean	3rd quartile	Maximum
Maternal mortality ratio	9	271.2	393.5	431.7	600.5	1360.0
Current health expenditure	2.531	4.301	5.441	6.055	7.053	18.319
Gross domestic product	304.4	652.9	1225.8	2355.5	2963.2	14745.3
Human development index	0.2850	0.4452	0.5105	0.5325	0.5910	0.7910
Health access and quality index	28.60	41.67	45.05	47.37	50.85	70.10
Antenatal care(4visits)	6.30	45.65	56.60	56.96	75.05	87.30
Total fertility rate	1.360	3.524	4.620	4.358	5.046	7.290
Female Literacy rate	13.93	33.58	62.44	57.52	76.81	94.45
Physician density	0.019	0.055	0.095	0.312	0.395	2.092
Births by skilled	19.40	53.40	72.20	68.22	88.28	99.80

The median and the mean are good measures of central tendency used to describe the data. If the data are not normally distributed, the median is preferred rather than the mean but in the case of normally distributed data both measures are good to describe the data.

In the year 2011 the median was 428, meaning that in 2011 half of African countries had maternal mortality ratios less than 428 deaths per 100 000 live births and the other half had more than that. In 2012 the median of 415.5 indicates that half of the countries had maternal mortality less than about 416 deaths per 100 000 live births and the other half had greater ratios. For the subsequent years; 2013, 2014 and 2015 the median values are 411.5,400.5 and 393.5 respectively. This shows that maternal mortality has been decreasing throughout the years.

In 2011 the lowest and the highest ratios were 9 and 1580 deaths per 100 000 live births respectively. Maternal mortality has decreased over time from 2011 and led to the maximum value of maternal mortality ratio being 1360 deaths per 100 000 in 2015.

The other variables will be used as the predictor variables in the study. Their descriptions are also interpreted in the same manner of identifying the minimum, maximum, mean and median to observe changes in years.

4.1.1 Kendall's Tau correlation results. Some of the variables under the study are found to follow a normal distribution and some follow a non-normal distribution. Kendall's Tau correlation test is used to

avoid making assumptions about the distributions of the data.

Table 4.7: **Kendall correlation coefficients between maternal mortality and the predictors**

	Kendall Tau correlation coefficient test					
Predictors	**Statistic**	**2011**	**2012**	**2013**	**2014**	**2015**
Current health expenditure	τ	-0.0559	-0.0832	-0.0538	-0.0294	-0.0322
	p-value	0.5556	0.37868	0.57072	0.75969	0.73708
Gross domestic product	τ	-0.4430	-0.4745	-0.4451	-0.4430	-0.4486
	p-value	2.3289e-06	4.2334e-07	2.087e-06	2.3289e-06	1.7334e-06
Human development index	τ	-0.6129	-0.6122	-0.5996	-0.5835	-0.5821
	p-value	6.3169e-11	6.6478e-11	1.615e-10	4.8877e-10	5.3633e-10
Health access and quality index	τ	-0.5220	-0.5227	-0.5395	-0.5318	-0.5276
	p-value	2.6007e-08	2.4939e-08	8.8126e-09	1.4183e-08	1.8401e-08
Antenatal care(4visits)	τ	-0.3140	-0.3150	-0.2960	-0.2860	-0.2690
	p-value	0.0011908	0.0011255	0.002257	0.0031916	0.0056034
Total fertility rate	τ	0.5702	0.5604	0.5430	0.5395	0.5381
	p-value	2.22e-16	2.22e-16	2.22e-16	2.22e-16	2.22e-16
Female literacy rate	τ	-0.5031	-0.5052	-0.5108	-0.5101	-0.5115
	p-value	8.1319e-08	7.1879e-08	5.1488e-08	5.3645e-08	4.9331e-08
Physician density	τ	-0.5779	-0.5912	-0.5884	-0.5625	-0.5639
	p-value	3.0116e-09	1.2315e-09	1.4904e-09	8.3097e-09	7.5865e-09
Births by skilled attendants	τ	-0.5420	-0.5395	-0.5395	-0.5122	-0.5087
	p-value	8.3707e-09	8.7959e-09	8.8126e-09	4.7073e-08	5.8095e-08

The correlation coefficients of current health expenditure, Gross domestic product, human development index, health access and quality index, antenatal care (4visits), female literacy rate, physician density and births attended by skilled personnel indicate negative relationships with maternal mortality throughout the years. Current health expenditure even though it has a negative relationship its effect is insignificant.

The strength of the relationship between gross domestic product and maternal mortality ratio is moderate throughout the years, (with -0.443 in 2011, -0.4745 in 2012, -0.4451 in 2013, -0.4430 in 2014, -0.4486 in 2015). The strength of the relationship between human development index and maternal mortality ratio is moderate throughout the study period with the coefficients between -0.5 and -0.6. Health access, physician density, female literacy rate and births attended by skilled personnel are also moderately related to maternal mortality with coefficients around -0.5. These variables are said to have a moderate negative relationship with maternal mortality. The strength of the relationship between percentage of women who attend at least four antenatal visits and maternal mortality is weak and we refer to the relationship as a weak negative relationship.

When gross domestic product per capita increases maternal mortality ratio decreases moderately. When the percentage of women who attend at least four antenatal visits increases maternal mortality ratio slightly decreases. For the rest of the other variables negatively correlated with maternal mortality ratio, when they increase maternal mortality decreases depending on the strength of the correlation.

Total fertility rate is positively correlated with maternal mortality for all the years. Correlation coefficients of total fertility rate for all the years under the study are around 0.5 indicating a moderate strength of relationship. Total fertility rate has a moderate positive relationship with maternal mortality, when total fertility rate increases maternal mortality also increases.

4.2 Mapping results

For the five years under the study the pattern of maternal mortality has been seen to be similar for all years. Figure 4.1 describes only the 2015 (latest available) data on Maternal mortality ratio of African countries. Countries with high maternal mortality are coloured with the deepest colour and gets lighter while maternal mortality decreases, low maternal mortality countries are then coloured in a lightest colour.

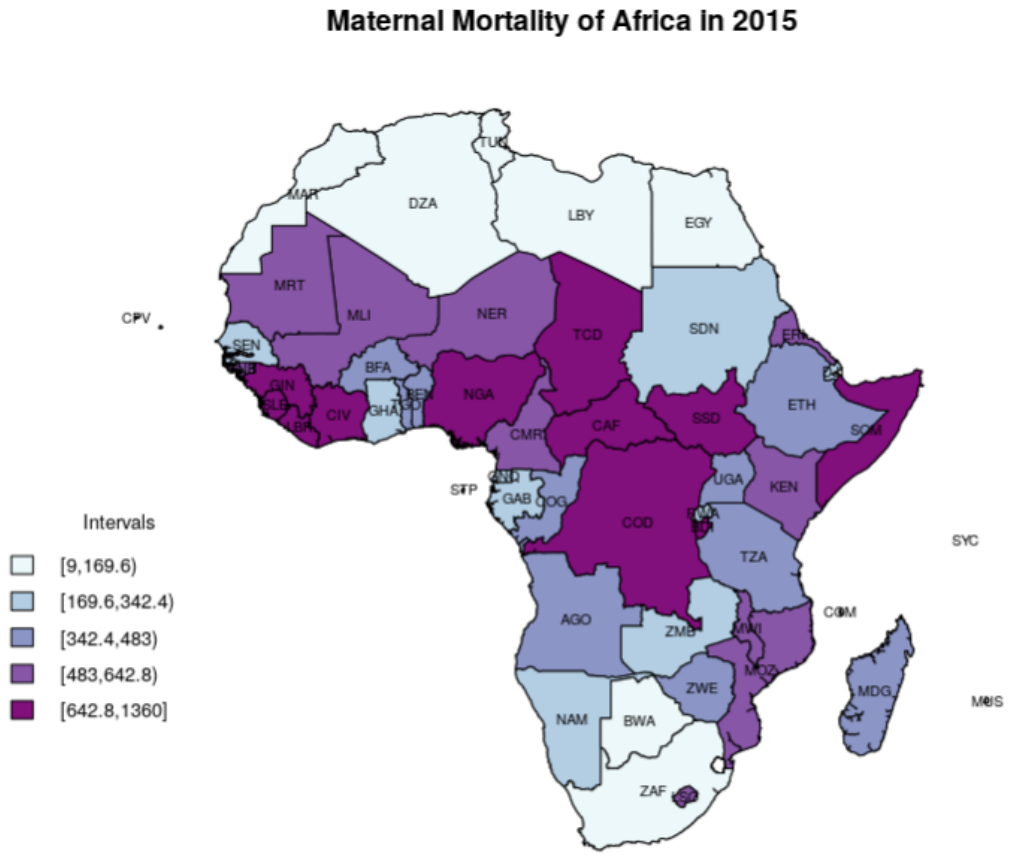


Figure 4.1: **Map of Africa showing incidence of maternal mortality by country**

The map tells us that Maternal mortality is low in countries like Algeria, Botswana, Egypt, Morocco, Libya, Namibia, South Africa and Zambia. Countries like Congo DRC, Chad, South Sudan, Nigeria, Somalia, Cote d'Ivoire, Sierra Leone, Guinea, Liberia and Central African Republic have high Maternal mortality ratios. It is observable that most of the countries with low maternal mortality ratios are in the northern part of the continent, then some in the southern of the continent. The region with most of the countries having high maternal mortality ratios is West Africa.

4.3 Moran's I results

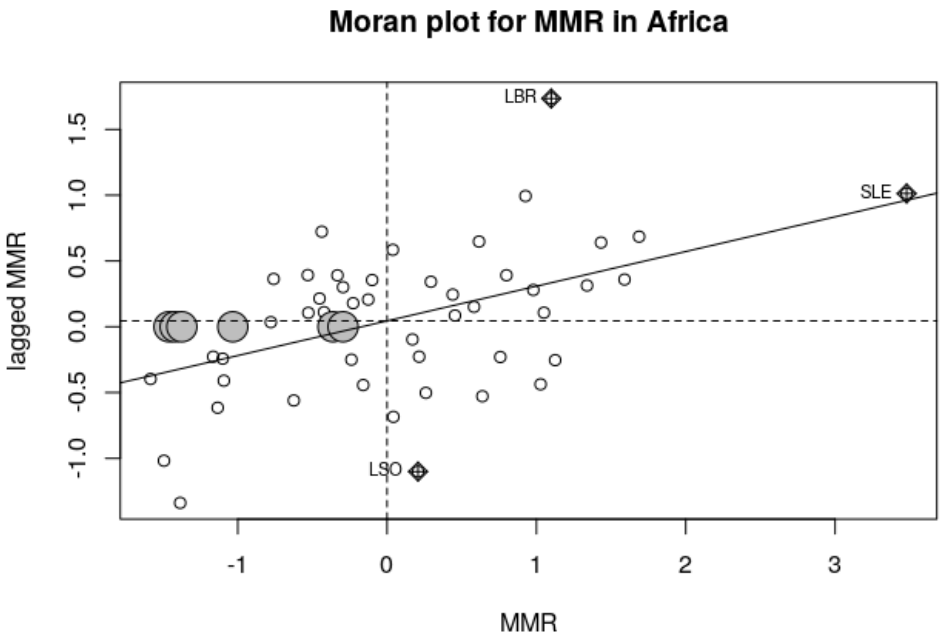
Table 4.8: **Moran's I test statistics under randomization assumption by years**

Statistic / Year	Moran's I	Expected	std deviate (Z_I)	Variance	p-value
2011	0.263178569	-0.021276596	2.9611	0.009228493	0.001533
2012	0.267226794	-0.021276596	2.9909	0.009304752	0.001391
2013	0.266216533	-0.021276596	2.9741	0.009344513	0.001469
2014	0.263546170	-0.021276596	2.9407	0.009380969	0.001637
2015	0.263607333	-0.021276596	2.9364	0.009412449	0.00166

Table 4.8 shows the Moran's I test results of maternal mortality for the years 2011-2015. For the year 2015 the Moran's I statistic (0.2636) is positive and it is greater than the expected value meaning there is a positive autocorrelation. We can also use the p-value to conclude about the hypothesis. The p-value (0.00166) is significant as a result we reject the null hypothesis and say that maternal mortality in Africa is positively autocorrelated. When we have positive spatial autocorrelation we conclude that ratios of maternal mortality in Africa tend to be similar for neighbouring countries, thus we say they form a spatially clustered pattern.

For all the other four years 2011-2014 the estimated values are slightly the same as those obtained when testing for 2015 maternal mortality ratios. We then notice that for all these other years we have the same pattern; positive spatial autocorrelation thus we say the spatial pattern of maternal mortality is also clustered just like in 2015.

Figure 4.2: **The Moran's I scatter plot of Maternal mortality in 2015**



The scatter plot in Figure 4.2 shows the spatial autocorrelation in Africa. The slope is the Moran's I index showing a positive autocorrelation. Labelled countries are those that their maternal mortality have high influence in the analysis. Labelled countries are in the high-high and low-low quartiles which also indicate a positive autocorrelation of African countries.

The plot shows that Liberia and Sierra Leone are in the high-high quadrant, they have high maternal mortality and they are surrounded by countries with high maternal mortality. Lesotho shows up in the high-low quadrant, this indicates that Lesotho has high maternal mortality ratio but its neighbours have low maternal mortality ratios. Lesotho can learn from neighbouring countries how the reduce maternal mortality.

Neighbouring countries have similarities sometimes because they are exposed to similar factors that may affect Maternal mortality. Below we focus on spatial autocorrelation within regions to see if the ratio of maternal mortality may depend on neighbouring countries belonging to the same region.

Table 4.9: **Moran's I test statistics under randomization assumption**

Statistic / Region	Moran's I	Expected	std deviate (Z_I)	Variance	p-value
East Africa	-0.11379653	-0.11111111	-0.013518	0.03946199	0.5054
West Africa	0.17796572	-0.06250000	1.5763	0.02327299	0.05748
Central Africa	0.3179393	-0.2500000	1.6338	0.1208324	0.05115
North Africa	-0.3944076	-0.2000000	-0.82244	0.0558754	0.7946
Southern Africa	-0.10431751	-0.11111111	0.035425	0.03677723	0.4859

Maternal mortality ratios for the years 2011-2015 are discovered to be relatively the same. In Table 4.9 maternal mortality ratio data by regions only for the latest year (2015) are considered.

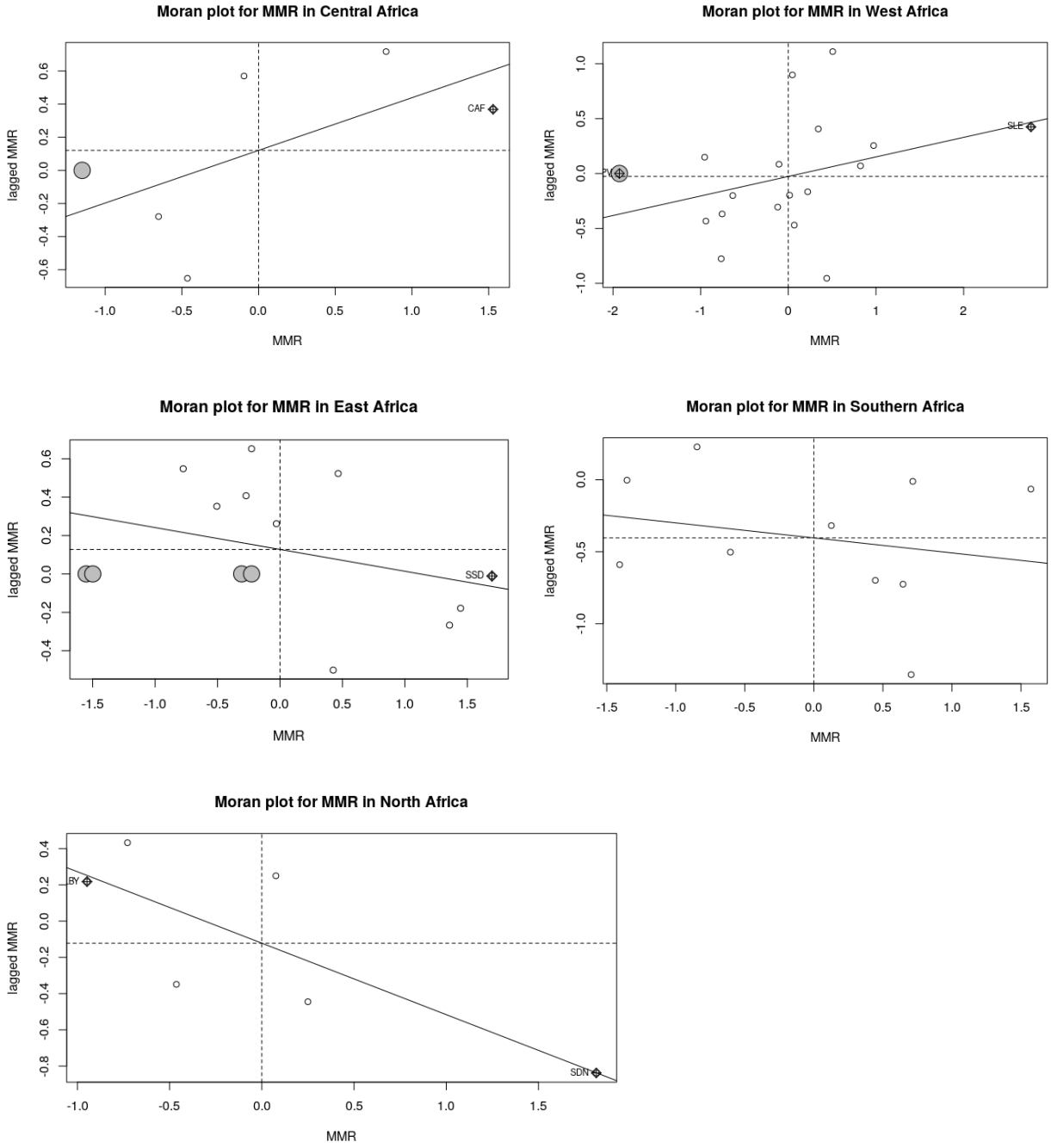
In East Africa, the Moran's I is slightly less than zero, it shows a negative spatial autocorrelation and that countries in the East are heterogeneous. This means the maternal mortality ratio of one country in East Africa does not depend significantly on values of maternal mortality ratio of other countries in East Africa, the ratios are not similar. This dissimilarity is not so important because the associated p-value is not significant.

North Africa also shows that there are countries that are negatively autocorrelated. The maternal mortality patterns are not similar for all countries in the north. Looking at the statistics of the southern part of Africa and East Africa negative spatial autocorrelations are also observed. The spatial patterns of the East, North and Southern Africa are referred to as dispersed patterns. Maternal mortality ratios are not similar regardless of whether they are neighbours and all belong to the same region. In these regions countries with high maternal mortality can learn from their neighbouring countries doing well. But the spatial negative autocorrelation for these regions do not have a big effect in the analysis, they are shown to be insignificant (p-value$>$0.05).

West and Central Africa show more or less the same results and their results mean the same thing. There is a low level of positive spatial autocorrelation, some of the countries per region have similar maternal mortality ratios. The patterns in these regions are clustered and all have a significant autocorrelation. Countries with high ratios of maternal mortality are surrounded by countries with high ratios or the case could be that countries of low maternal mortality ratios are among those with low maternal mortality too.

Scatter plots in Figure 4.3 below are used to describe the autocorrelations of maternal mortality.

Figure 4.3: **Moran scatter plots of Maternal mortality in African regions in 2015**



The negative slopes of the last three plots in the figure above indicate that in the East, North and Southern Africa the spatial autocorrelation is negative as explained by Moran's I value in Table 4.9. In East Africa we see that South Sudan has high maternal mortality ratio but surrounded by countries with low maternal mortality ratios. The North Africa plot shows that in North Africa, Libya has a low maternal mortality ratio but surrounded by countries with high maternal mortality. Sudan is one of

those countries with high maternal mortality ratio surrounded by countries like Libya with low ratios. In Southern Africa, countries are negatively autocorrelated but the autocorrelation does not specify which countries had huge influence in the analysis.

As indicated by the Moran's I in Table 4.9, the slopes of the scatter plots for Central Africa and West Africa show a positive spatial autocorrelation. In Central Africa, Central African Republic has high maternal mortality and its neighbours also have high maternal mortality above the mean. In West Africa Sierra Leone with a high maternal mortality has its neighbouring countries with high maternal mortality too. Cape Verde has a ratio close to mean and it is surrounded by neighbours with ratios not so far away from the average value of maternal mortality.

Generally, maternal mortality may or may not depend on the location of countries. There are countries that do not have similar ratios of maternal mortality regardless of weather they belong to the same region or not.

4.4 Panel fractional regression results

4.4.1 Hausman test results.

$$\chi^2_9 = H = 32.576 \text{ and; } p - value = 0.0001583$$

The output shows that the p-value is significant at 5% level of significance. The fixed effects estimator is more efficient than the random effects estimator. Even though the fixed effects model is preferred over random effects Schunck (2013) states that it has a draw back. a model that estimates within effects in random models is better. A fixed effects model does not estimate effects of variables that do not change over time. Thus the correlated random effects estimator will be used in the study. The regression model used is the quasi maximum likelihood of correlated random effects using probit link function as proposed by Papke and Wooldridge (2008).

Table 4.10: **Panel fractional response regression output**

Variable	Estimate	Std.Error	t value	Pr(>|t|)	
Current health expenditure	-0.000049	0.002140	-0.023	0.982	
Gross domestic product	-0.000001	0.000002	-0.353	0.724	
Human development index	-0.659553	0.294349	-2.241	0.025	**
Total fertility rate	0.055913	0.046679	1.198	0.231	
Antenatal care(4visits)	-0.001057	0.001059	-0.997	0.319	
Health access and quality index	0.003307	0.003027	1.093	0.274	
Physician density	0.008669	0.040342	0.215	0.830	
Female literacy rate	0.000780	0.000838	0.930	0.352	
Births by skilled attendants	0.000244	0.000391	0.622	0.534	
Year.2012	-0.005528	0.002641	-2.093	0.036	**
Year.2013	-0.010237	0.006531	-1.567	0.117	
Year.2014	-0.014153	0.010679	-1.325	0.185	
Year.2015	-0.016208	0.013756	-1.178	0.239	
INTERCEPT.mean	-2.473343	0.275813	-8.967	0.000	***
Current health expenditure.mean	0.017320	0.012824	1.351	0.177	
Gross domestic product.mean	0.000011	0.000015	0.726	0.468	
Human development index.mean	0.393947	0.696546	0.566	0.572	
Total fertility rate.mean	-0.007232	0.050476	-0.143	0.886	
Antenatal care(4visits).mean	0.003244	0.001581	2.052	0.040	**
Health access and quality index.mean	-0.008651	0.005482	-1.578	0.115	
Physician density.mean	-0.278370	0.148196	-1.878	0.060	*
Female literacy rate.mean	-0.002443	0.001498	-1.631	0.103	
Births by skilled attendants.mean	-0.001722	0.001470	-1.172	0.241	

In the regression output there are other explanatory variables in the model that do not significantly influence the change in maternal mortality. We focus more on those that are statistically significant in the analysis. Statistically significant variables in the analysis are the year 2012, human development index, physicians density per 1000 people and percentage of women who visit antenatal care at least four times.

The time variables are shown to be insignificant except for the year 2012 with the year 2011 as the reference year. The analysis of maternal mortality during over time is done with reference to the year 2011. The year 2012 is significant with a p-value of 0.036 at 0.05 level of significance. The regression estimate for the year 2012 is -0.005528. This indicates that the maternal mortality ratio had a significant decrease in 2012. Over the other years the change in maternal mortality was not significant. The intercept mean coefficient shows that if all the predictors are held constant maternal mortality can be decrease by about 2.5 deaths per 100 000 live births.

From the regression results we observe that there is a significant negative relationship between human development index and maternal mortality at $\alpha = 0.05$. The regression estimate of -0.659553 suggests that a unit increase of human development index will decrease maternal mortality ratio by approximately 0.7 deaths per 100 000. Two units increase of human development index will decreases maternal mortality by about 1.4 deaths, meaning one life can be saved. If the human development index of a country is high, maternal mortality ratio is expected to be low.

On average physician density per 1000 people has a negative relationship with maternal mortality. The regression coefficient is -0.278370 with an associated p-value of 0.06 at ($\alpha = 0.1$). An increase by one medical physician per 1000 people will decrease maternal mortality by 0.28 deaths. If we have more physicians to attend to pregnant women it means maternal health care will be better thus reducing the ratio of maternal mortality.

On average, the percentage of women who attend antenatal care at least four times is significant at $\alpha = 0.05$ with an associated p-value of 0.040. It shows an unexpected result of having a positive relationship with maternal mortality ratio. It increases maternal mortality by 0.003244 deaths. This could be caused by countries whereby people attend antenatal care but the quality of the service is not good. If we look at just antenatal care coverage not on average we discover that its unit increase will decrease maternal mortality but the effect is insignificant, decreases maternal mortality by 0.001 deaths which is a very small change.

5. Conclusion and Recommendations

Maternal mortality ratios of countries can be similar sometimes because they are closer in space and thus may be exposed to similar factors that influence maternal mortality. For example, Liberia and Sierra Leone are western countries both with high maternal mortality. However, there are countries that are not close in space but have significant similarities of factors that influence maternal mortality. Central African Republic and Sierra Leone both have high maternal mortality but belong to different regions. On the contrary, some countries can still be in one region but have disparities in maternal mortality because of different levels of associated factors.For instance, Libya has low maternal mortality ratio and Sudan has high maternal mortality ratio but they both belong to North Africa. It was also observed that maternal mortality has declined from 2011-2015 but not satisfactorily.

Under bivariate analysis, maternal mortality ratio per 100,000 live births was found to have a significant inverse relationship with gross domestic product per capita, human development index, percentage of women who attend antenatal care at least four times, health access and quality index, physicians per 1000 people, female literacy rate and births attended by skilled health staff. These variable need to increase to improve maternal mortality in Africa. Only total fertility rate had a direct relationship with maternal mortality ratio. Women should be advised not to have many children as it increases chances of dying.

From the multivariate analysis human development index, physician density and percentage of women who attend antenatal care at least four times are found to be significantly related with maternal mortality. Human development index assesses the development of the country economically and socially. A good level of Human development index indicates that the levels of education, per capita income and long healthy life are high. Availability of sufficient medical physicians needs to increase in order to fight high ratios of maternal mortality. To reduce maternal mortality; health, education and the living standard of people need to be improved. The unexpected results of how the percentage of women attending at least four antenatal visits affects mortality may be caused by the fact that in other countries, even if women attend antenatal care services the care is not of good quality.

Most importantly the healthcare system needs attention, people might have access and enough physicians but not receive the appropriate quality of antenatal care. There should be an intervention whereby countries with better health systems visit other countries to work there to share some working strategies. People because of ignorance do not take good care of themselves and do not understand the importance of attending health care services. Policies for 'education for all' should be implemented in all African countries because education seems to have a bad impact on maternal mortality if it is low.

The latest available data on the variable of interest (maternal mortality ratio) is of the year 2015, that is four years ago. Maternal mortality can be affected by many other indicators not included in the study, but because of missing information on many countries they had to be omitted. Absence of data in Africa is a concern, data needs to be accurately collected, reported and made available to researchers.

Acknowledgements

A special thanks I give to my supervisor Dr. Siaka Lougue who has guided me through this essay. I am grateful for his precious time and being available to me. I would also like to extend my gratitude to the African Institute for Mathematical Sciences (AIMS), The academic directors, my lectures and tutors who had been a great source of inspiration. I would also like to thank AIMS staff members and my colleagues who were like brothers and sisters to me.

I place a deep sense of appreciation to Mr. Tlokotsi Mpeka for his unfailing love and support that kept me going throughout my studies. I also wish to express my thanks and appreciation to my family, friends and everyone who provided me the possibility to complete this essay.

Above all, I thank the Lord Almighty for his guidance and granting me the strength and knowledge to undertake this study.

References

Kendall's tau-b correlation tests (simulation). https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/PASS/Kendalls_Tau-b_Correlation_Tests-Simulation.pdf, Accessed May 2019.

Alimohamadi, Y., Khodamoradi, F., Khoramdad, M., Shahbaz, M., and Esmaeilzadeh, F. Human development index, maternal mortality rate and under 5 years mortality rate in west asian countries, 1980-2010. *Eastern Mediterranean Health Organization Journal*, 2017.

Alvarez, J. L., Gil, R., Hernández, V., and Gil, A. Factors associated with maternal mortality in sub-saharan africa: an ecological study. *BMC Public Health*, 2009.

Bayati, M., Vahedi, S., Esmaeilzadeh, F., Kavosi, Z., Jamali, Z., Rajabi, A., and Alimohamadi, Y. Determinants of maternal mortality in eastern mediterranean region: A panel data analysis. *Medical Journal of the Islamic Republic of Iran*, 30:360, 2016.

Berhan, Y. and Berhan, A. Skilled attended delivery health personnel as a proxy indicator for maternal and perinatal mortality. *Ethopian Journal of Health Sciences*, 24:69–80, 2014.

Calabrese, R. Regression Model for Proportions with Probability Masses at Zero and One. Working Papers 201209, Geary Institute, University College Dublin, Mar. 2012.

Diorio, A. and Dr.Crivelli-Kovach, A. The global and local factors influencing maternal mortality ratios: Barriers and recommendations for success. *Journal of Social Science for Policy Implications*, 2(3): 33–60, 2014.

Ensor, T., Cooper, S., Davidson, L., Fitzmaurice, A., and Graham, W. J. The impact of economic recession on maternal and infant mortality: lessons from history. *BMC Public Health*, 10(1):727, 2010.

Gallani, S., Krishnan, R., and Wooldridge, J. Applications of fractional response model to the study of bounded dependent variables in accounting research. *SSRN Electronic Journal*, 01 2015.

Geary, R. C. The contiguity ratio and statistical mapping. *The Incorporated Statistician Journal*, 5(3): 115–146, 1954.

Girum, T. and Wasie, A. Correlates of maternal mortality in developing countries: an ecological study in 82 countries. *Maternal Health, Neonatology and Perinatology Journal*, 2017.

Hausman, J. A. and Leonard, G. K. Superstars in the national basketball association: Economic value and policy. *Journal of Labor Economics*, 15(4):586–624, 1997.

Hossain, S. Panel data. fixed effect and random effect. model two. eviews. https://www.youtube.com/watch?v=BkhyBDSeq68, Accessed May 2019, 2013.

Johnson, C. The truth about maternal death. The Conversation, https://theconversation.com/the-truth-about-maternal-death-94361, Accessed March 2019.

Koski-Rahikkala, H., Pouta, A., Pietiläinen, K., and Hartikainen, A.-L. Does parity affect mortality among parous women? *Journal of Epidemiology and Community Health*, 60(11):968–973, 2006.

Lee, D., Dean, N., and Anderson, C. Identifying clusters in Bayesian disease mapping. *Journal of Biostatistics*, 15(3):457–469, 03 2014.

Liu, W. and Xin, J. Modeling fractional outcomes with sas. *SAS Support Resource*, 2014.

McCarthy, J. and Maine, D. A framework for analyzing the determinants of maternal mortality. *Journal of Studies in Family Planning*, 23(1):23–33, 1992.

Meh, C. *Determinants of Maternal Mortality: A Comparison of Geographic Differences in the Northern and Southern Regions of Cameroon*. Masters, The University of Western Ontario, 2017.

Murray, D. Maternal mortality rate, causes and prevention. https://www.verywellfamily.com/maternal-mortality-rate-causes-and-pevention, Accessed March 2019, 2019.

National Academies Press, 2000. Evidence on the consequences of maternal mortality. https://www.ncbi.nlm.nih.gov/books/NBK225432/?report=reader, Accessed April 2019.

Papke, L. E. and Wooldridge, J. M. Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of Econometrics*, 11(1):619–632, 1996.

Papke, L. E. and Wooldridge, J. M. Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics*, 145(1):121 – 133, 2008.

Ramalho, E. A., Ramalho, J. J., and Coelho, L. M. Exponential regression of fractional-response fixed-effects models with an application to firm capital structure. *Journal of Econometric Methods*, 7(1), 2016.

Rana, R. H., Alam, K., and Gow, J. Health expenditure, child and maternal mortality nexus: a comparative global analysis. *BMC International Health and Human Rights*, 18(1):29, Jul 2018.

Rezaeian, M., Dunn, G., St.Leger, S., and Appleby, L. Geographical epidemiology, spatial analysis and geographical information systems: a multidisciplinary glossary. *J Epidemilogy Community Health*, 61 (2):98–102, Feb 2007.

Schunck, R. Within and between estimates in random-effects models: Advantages and drawbacks of correlated random effects and hybrid models. *The Stata Journal*, 13(1):65–76, 2013.

Shapiro, S. S. and Wilk, M. B. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4):591–611, 1965.

Torres-Reyna, O. Panel Data Analysis Fixed and Random Effects using Stata . *Data and statistical services*, 4(2), Dec 2007.

UNICEF, 2017. Maternal mortality. https://data.unicef.org/topic/maternal-health/maternal-mortality/, Accessed March 2019.

UNICEF, 2018. Antenatal care. https://data.unicef.org/topic/maternal-health/antenatal-care/, Accessed March 2019.

WHO, 2007. *Maternal mortality in 2005: estimates developed by WHO,UNICEF, UNFPA, and the World Bank.* 2007.

WHO, 2013. *Maternal Death Surveillance and Response: Technical guidance information for action to prevent maternal death*. 2013.

WHO, 2015. *Trends in maternal mortality: 1990-2015*. World health organisation.

WHO, 2018. Maternal mortality. https://www.who.int/news-room/fact-sheets/detail/maternal-mortality, Accessed March 2019.

Ye, J., Zhang, J., Mikolajczyk, R., Torloni, M., Gulmezoglu, A., and AP, B. Association between rates of caesarean section and maternal and neonatal mortality in the 21st century: a worldwide population-based ecological study with longitudinal data. *Journal of Social Science for policy Implications*, 123, Aug 2015.