ADDIS ABABA UNIVERSITY

OFFICE OF GRADUATE PROGRAMME

DEPARTMENT OF STATISTICS

Determinants of Migration to South Africa Using Count Regression Model: The Case of Soro and Misha Districts in SNNPR

Sebsibe Admasu

Advisor: Shibru Temesgen (Dr.)

A Thesis Submitted to the department of Statistics of Addis Ababa University in Partial Fulfillment of the Requirements for the Degree of Master of Science in Statistics (Applied Statistics)

June, 2018
Addis Ababa University
Addis Ababa, Ethiopia

ADDIS ABABA UNIVERSITY

OFFICE OF GRADUATE PROGRAMME

DEPARTMENT OF STATISTICS

Determinants of Migration to South Africa Using Count Regression Model: The Case of Soro and Misha Districts in SNNPR

Sebsibe Admasu

A Thesis Submitted to the department of Statistics of Addis Ababa University in Partial Fulfillment of the Requirements for the Degree of Master of Science in Statistics (Applied Statistics)

Advisor: Shibru Temesgen (Dr.)

June, 2018
Addis Ababa University
Addis Ababa, Ethiopia

Addis Ababa University

School of Graduate Studies

This is to certify that the Thesis prepared by Sebsibe Admasu, entitled: for Determinants of Migration to South Africa Using Count Regression Model: The Case of Soro and Misha District in SNNPR and submitted in partial fulfillment of the requirements for the Degree Master of Science in Statistics complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the Examining Committee:

Name	Signature	Date
Advisor: Shibru Temesgen (Dr.)	.....	.....
Examiner: Mekonnen T. (Assoc. Prof.)	.....	.....
Examiner: Dejen Tesfaw (Dr.)	.....	.....

DECLARATION

I, the undersigned, declare that the thesis is my original work, has not been presented for degrees in any other University and all sources of materials used for the thesis have been duly acknowledged.

Name: Sebsibe Admasu

Signature: _____

Place of submission: Department of statistics, College of Natural and computational sciences, Addis Ababa University

Date of submission: June, 2018

This thesis has been submitted for examination with my approval as a University advisor.

Shibru Temesgen (D.r)_____

Advisor's Name **Signature**

Acknowledgments

Thanks to ***JESUS*!** This thesis could not have been accomplished without assistance of a number of persons and I would like to take these few lines to thank them warmly for their help. Firstly, I extend my heartfelt appreciation and thanks to my **Advisor Dr. Shibru Temesgen** for the guidance rendered to me in the doing of this Thesis. I also thank for taking your time to advice through my work, as well as the criticisms and the suggestions. It is a real pleasure working with him. I can only hope that his cooperation will keep on going in the future.

Secondly, I am grateful to the entire staff of **Statistics Department of Addis Ababa University** for their support by various inputs. My stay in the university was also a great pleasure. Not only did I enjoy the past two years that I have spent here, but I also learned a lot through many short term trainings.

Finally, I am also very grateful to all persons whose presence and support have always been important to me.My thanks go to my **classmates** and **friends**. My heart-felt thanks go to my father **Admasu Songaye**, my Mother **kebebush Gintamo**, my brothers **Zerihun**, **Gashaye** and **Bereket** and my sister **Melesech** for their interest for this Thesis from the beginning up to this stage by giving a great support by finance, suggestion and materials. I have said may God bless you. Thank you once again for your kind generosity.

Contents Of Table

List of Tables

List of Figures

ACRONYMS

AIC: Akaike information criterion

BIC: Bayesian information criterion

CSA: Central Statistical Agency

GDP: Gross domestic product

HOA: Horn of Africa

HH: Household

HHH: Household head

IOM: International Organization of Migration

IRR: Incidence Rate Ratio

LRT: Likelihood-ratio test

MCMC: Markov Chain Monte Carlo

ML: Maximum Likelihood

NB: Negative binomial

OLS: Ordinary Least Squares

Pmf: Probability mass function

PPS: Population proportional allocation

RSA: Republic of South Africa

SE: Standard error

SNNPR: Southern nation nationality regional government

UNDESA: United Nations Departments of Economic and Social Affairs,

UN: United nation

U.S: United State

ZI: Zero-Inflated

ZIP: Zero-Inflated Poisson

ZINB: Zero-Inflated Negative Binomial

Abstracts

Determinants of Migration to South Africa Using Count Regression Model: The Case of Soro and Misha Districts in SNNPR

Sebsibe Admasu

Addis Ababa University, 2018

The labor power migration affects the socioeconomic development of the people in developing countries like Ethiopia. The goal of this study was to explore the determinants that affect the number of migrants per household in Soro and Misha woreda, Hadiya zone and to investigate impacts of migration in the Society which is related to the causes and consequences of labor power migration. To achieve these objectives, both quantitative and qualitative methods were employed, we used mainly primary data and for more information secondary data also were used for particular purpose. Primary information was collected mainly from the households and secondary data was collected from Labour and Social Affairs Bureaus. The data were gathered from randomly selected kebeles from the Soro and Misha District. The tool which was used to gather the primary information was questionnaires and interviews. Total 754 respondent households were selected for survey questionnaire by stratified sampling technique. Descriptive statistical method was employed to analyze quantitative data by using STATA and excel whereas to explore the determinant factors of migration, Count Regression model was applied by using R and STATA software. Among the count models considered, the Zero-Inflated Poisson (ZIP) model was found to be the most appropriate model for analyzing the number of migrants. The ZIP model results indicated that religion, HHH age, HHH education, place of residence, land size, family size, dependency ratio and HHH job occupation are significantly associated with the number of migrants per households. The reveals that lack of job opportunities, family or peer pressure, poverty, and unemployment were identified by this study as the key push factors of migration. On the other hand, job opportunities, better income, social networks and smugglers at destination country were considered as pull factors of migration. Regarding the problem on the journey imprisonment, robbery, lack of food and water, and human trafficking were identified.

CHAPTER ONE

1 INTRODUCTION

1.1 Background of the Study

Migration is relatively permanent change of residence or address involving crossing an administrative or political boundary. It is mainly classified into two types: internal and international migration. Internal migration is defined as, the change of the place of residence from one administrative border line to another within the same country, while international Migration is a movement in excess of a national border line. Migration is the movement of a person or a group of persons, either across an international border, or within a State. It is a population movement, encompassing any kind of movement of people, whatever its length, composition and causes; it includes migration of refugees, displaced persons, economic migrants, and persons moving for other purposes, including family reunification (IOM, 2014).

In most discussions on migration, the starting point is usually numbers. The current global estimate is that there were around 244 million international migrants in the world in 2015, which equates to 3.3 percent of the global population (UNDESA, 2016).A first important point to note is that this is a very small minority of the global population, meaning that remaining within one's country of birth overwhelmingly remains the norm. The great majority of people in the world do not migrate across borders; much larger numbers migrate within countries (an estimated 740 million internal migrants in 2009 (UNDP, 2009).

In recent years we have also seen a significant increase in displacement, both internal and across borders, which has largely stemmed from civil and transnational conflict including acts of violent extremism outside actual warzones. Current data indicate that in 2016 there were 40.3 million internally displaced persons (IDPs) world wide and 22.5 million refugees (IUNHCR, 2017)..Importantly, the beneficial effects of migration for migrants and their families go beyond economic impacts and frequently include improvements in other dimensions of human development, such as education and health.

For example, according to a recent report by the World Bank, "migrants from the poorest countries, on average, experienced a 15-fold increase in income, a doubling of school enrolment rates and a 16-fold reduction in child mortality after moving to a developed country (World Bank 2016). One tends to ignore the fact that the developing world is not homogenous that some states are more developed than others. As such, the relatively more developed states in the developing world experience many of the same problems that more developed counterparts in the Western world experience.

Today, we have to agree without any hesitation that international migration became a crucial part of our global life. Like trade, finance, climate change and environmental issues, migration has got unprecedented global dimensions. Specially, since the end of the cold war, migration has become a factor in a relationship between governments because of its significant direct and indirect impacts on the economies, demographics, and daily life in countries of origin, transit and destination (World Bank, 2016).

The determinant factors of international migration are diverse and complex. From the reasons classified as "push and pull" factors, globalization imposed the implementation of a third set of motivations under the designation of "network", which includes the free flow of information, improvement of communication at world level, along with swifter and cheaper transport. These latter factors are not a direct migration reason, but elements contributing to favoring migration. It should be noticed that modern globalization is not determined only by the technological progress, but also by the political and ideological changes(Migration and Globalization,2010)The negative facts, such as economic disadvantages (unemployment, poverty), political or religious prosecution, and environmental effects of climate change, wars or natural calamity are included into push factors, positive facts, such as high level of social and economic life, freedom of political and religious thoughts and labor demand in capitalist markets should be considered as pull factors. International migratory movements in Africa have become more complex and mixed in recent years, with owes comprising asylum seekers, refugees and irregular migrants. The exodus of migrants from the Horn of Africa (mainly Ethiopia) to South Africa is a central issue. Each year, thousands of mainly young Ethiopians risk their lives in an attempt to reach South Africa, where they hope to establish better lives for themselves and their families.

1.2. Statement of the problem

The movement of people from place to place is an important component of population change which has its own spatial as well as temporal characteristics. In the era of globalization, an improved and sophisticated means of transportation and communication are playing a key role in facilitating the interaction of people around the world. As a result, the rate of migration is alarmingly increasing both at national and international level. In developing countries, internal migration is more persistent. In Ethiopia, migration from place to place is quite common especially in areas where drought is frequent (IOM, 2014).

Several things have been said about the migration of Ethiopian females to the Middle East countries (Abdu, 2009; Girum, 2010), but very little about the irregular migration of Ethiopians to the Republic of South Africa (RSA).Most of the young adults who migrate to RSA are economically active and are heading in pursuit of dream of capturing the green pasture there. In Ethiopia, the problem is widely observed in two zones of the southern parts, namely in Hadiya and Kambata-Tambaro Zones (Messay, 2005, Sinedu, 2009). The present study focuses on investigating the socioeconomic and demographic causes and consequences of illegal migration of people from Hadiya areas of southern Ethiopia to Republic of South Africa. This study is mainly aimed to explore the socioeconomic and demographic causes and consequences of migration of peoples from southern Ethiopia particularly from Hadiya zone Soro and Misha district to the Republic of South Africa. The reasons why this study is concerned with Soro and Misha districts is high proportion of international migrants has been reported by some literatures about these districts and high migration rate has been reported by Hadiya zone social security affairs office. And also these districts have always been affected by drought and shortage of rain fall as reported by Hadiya zone government communication office. With these facts in mind it is worthwhile to undertake this study as it is hoped that expected result of the analysis could be used for further related issues. In doing so, it tries to address the following research questions:

- What factors initiate peoples to migrate to Republic of South Africa?
- What are the socioeconomic and demographic consequences of this migration on households and the community at large?

1.3 Objective of the Study

1.3.1 General Objective

The objective of this study is to determine demographic and socioeconomic factors associated with number of migrants to RSA per household from Soro and Misha district in Hadiya zone, SNNPR.

1.3.2 Specific Objective of the Study

- To determine the effects of Demographic and Socio-economic factors for migration to RSA in the study area;
- To identify the Motivating pressures that make the society to migrate("pull" and "push" factors);
- To examine positive and negative impacts of migration to RSA in the environments.

1.4 Significance of the study

The findings of this research has relevance in producing information to the government of Ethiopia about the major factors that motivate the labor power migration and impacts of Migration from the Soro and Misha district to the Republic of South Africa (RSA). In addition, it is essential to identify the socio economic impacts of migration, and to contribute the appropriate measures to be taken on the negative consequences of migration in the area. To provide information that would be helpful for decision makers in formulating policies to mitigate the problems of illegal or irregular migration dominantly grown at study area. Finally, it can be used as a spring board for further studies on the issue raised.

1.5 Limitation of the study

Although many factors are affecting the number of migrant, this study is undertaken to explore a few socio-demographic factor that affect number of migrants per household. The following are some of the limitations of our study.

- The study was limited to capture all characteristics of environment due to lack of detailed satellite-based environmental characteristics in this study area and difficulties of deep observation by the researcher at limited budget and time.
- The study did not include all characteristics of respondent household.

To put off the defect due to these, researcher selected the relevant variables which are assumed to be significant and common environmental characteristic in affecting the livelihood of household and excludes the characteristics that are assumed to insignificant and highly correlated particularly in this study area

1.6. Organization of the paper

This thesis has been organized as follows. In Chapter one, the background, statement of the problem, objectives, scope and limitation of the study are presented. The second Chapter provides the literature review. The data source, methodology and the research hypotheses are given in Chapter three. The fourth Chapter is the analysis, findings and discussion Chapter. Finally, conclusions and recommendations are made in Chapter five.

CHAPTER TWO

2 .REVIEW OF LITRTURES AND CONCEPTUAL FRAMEWORKS

2.1THEORETICAL REVIEVIEW OF LITERATURE

2.1.1 DEFINITION AND CONCEPT OF MIGRATION

Migration is not a new phenomenon. Migratory movements are as old as mankind, as people crossed borders in search of a better life or livelihood, or in the rush to flee a failed State or natural disaster. Migration, which is population mobility within and outside given geographical boundaries has been almost always occurred in the past, and occurs regularly in the present times as well. It causes numerous impacts upon the social and economic lives of the migrants as well as their families Lee (1997)

Broadly speaking, migration can be categorized into two classes. Internal Migration: where people move within a given social system; and international migration where people cross the internationally recognized geographical and political boundaries, which separate one country from the other Lee (1997).

2.1.2 Push and pull factors

There are a number of theoretical models for explaining migration. Most notable among these is the “push-pull” theory propounded by Lee (1997). According to this theory there are certain "push" factors, which compel people to move out of their place of origin as a result of either bad climate, poverty, famine, war or religious and political conflicts or instability, etc. On the other hand, "pull" factors attract people to move out in their search for new opportunities or for a better life. These include better health, political and/or religious freedom, urban glamour and enjoyment, better education facility, improved medical facility, security, peaceful political atmosphere, religious tolerance and civic society Lee (1997).

Another explanation, provided by the human capital model, posits that the individuals’ decisions to move depend on the hope of being better off at the place of destination in relation to the existing circumstances at the place of origin. The individual compares the potential future benefits with the current costs of migration, and if the balance tilts in favors of the former, the individual decides to migrate (Davanzo, 1980). In other words, migration is considered as a human capital investment, which will yield net future gains. Todero (1976) has expressed similar

views also. He considers age, length of residence, occupation, and distance as the main factors contributing in the decision to migrate. If the net costs of these factors are less than the expected future returns from migration, people decide to move. (Anon 1996)

2.1.3 Voluntary and Forced Migration

Migration refers to a change of the usual place of dwelling (Arsole et al, 2003 in Adamnesh, 2008). It could be either voluntary or forced (Kokpari, 2000). Voluntary migrants are those who change places within or beyond their country of origin at their own discretion rather than for other uncontrollable factors while forced migration refers to the coerced movement of people away from their home or place of dwelling. People leave their place of origin permanently or temporarily to take advantage of opportunities in host countries. These opportunities could be economic, social, political, environmental or a combination of all. The tempting wage gaps between developed and developing countries, inviting immigration programs in the developed world, lack of democracy and good governance in the home countries, and poverty and environmental degradation in the home countries are some of the factors for international migration (Portes,1996).

2.1.4 Cumulative Causation theory

Cumulative causation theory identifies not one but several factors causing migration. Due to the fact that each act of migration appears unique and different from another, the social context within which the decision to migrate is made is also different. Thus the scholars applying this approach in migration studies identify the following elements which are affected by international migration.

2.1.5 Irregular Migration

The International Organization of Migration (IOM) defines irregular migration as a "migration that takes place outside the norms and procedures established by States to manage the orderly flow of migrants into, through, and out of their territories". An irregular immigrant is, thus, a person who infringes these norms and procedures. There are several categories of irregular migrants: people which enter/leave the country illegally and then stay even after visa expiry; either using false documents or in absence of any documents; people entering the country in unauthorized ways, but in this case against their will, being forced by organizations which are engaged in human trafficking and smuggling. Illegal entry is defined by the IOM as: "crossing

borders without complying with the requirements for legal entry into the receiving State". However, using the term "illegal" to define people is rather wrong, since the presence on the European territory might "not be authorized or their status as an immigrant may lack proper documentation, but it does not put them in a category where their very existence constitutes illegality".

2.2. Issues of Migration

In the migration literature, it is well established that in areas where migration is on its early stage, migration is expensive and risky and hence it is only the better-off households who will afford to send a migrant member abroad and benefit from remittances (Fajnzylber and Humberto Lopez, 2007). As a consequence, inequality will be aggravated while the effect on poverty will be insignificant. But, as migration increases through time, information about migration will be easily available which will make migration less costly and less risky. In the limit, even the very poor will afford to send migrant members abroad and benefit from remittances. At this stage, migration will be based on inequality and poverty reducing (Massey and Espinosa, 1997;Mckenzie and Rapoport, 2007).

One of the most obvious and direct effects of remittances is change in the distribution of income in the receiving country which in turn may affect poverty and inequality depending on who receives them. While the poor will have more demand to send some member(s) outside their community as a mechanism of improving their welfare. The fact that migration is usually expensive and risky also means that they are less likely to realize it in the absence of well functioning insurance and credit markets (Nwokocha, 2007).

Migration can be seen as an opportunity for households to spread risk by gaining access to incomes which are less affected by local constraints, thereby it can also be a way to overcome local market constraints in order to make investments in the local area (Taylor, 1999). However, these market constraints may also limit migrants' ability and incentive to invest in productive activities at their original home place (Taylor, 1999).

2.3. Migration in Ethiopia

Ethiopia had a population of 94millions in 2013(World Bank, 2014). The country was one of the top 10 migration corridors in 2010 and both a sender and a receiver of migration Ethiopia had an annual net migration rate of -0.23 migrants per 1,000 populations and an annual average net migration rate of 0.8 between2005 and 2010. In the same period, its stock of immigrants was estimated at 571,800 or 0.8% of the total population (UNDESA, 2013). Ethiopia's international migrant stock has not changed significantly since the 1990s. The percentage was 2.4% in 1990 and 1.0% in 2000. Between 2010 and 2013, the figure was slightly lower than 1 percent, at 0.7% and 0.8% respectively. The same period saw a significant shift in the annual rate of change in the migrant stock, from -5.6% in 1900-2000 to -1.3% in 2000-2010.

However, there was an even bigger change in 2010-2013, when the percentage leapt to about 7.1%. The stock of emigrants was estimated at 620,000 in 2010, representing 0.8% of the total population (World Bank, 2011, p.114) whereas female Ethiopian migrants constituted 46.2% of the international migrant stock in 2013 (UNDESA, 2013).

Sindu(2007) studied the international migration in Ethiopia was highly growing from the late 1970's due to the political instability in the country as the mentioned. Significant migration from Ethiopia to countries beyond the Horn of Africa (HOA) began after the 1974 Revolution. It is not denied that Ethiopia has experienced political instability, war, famine, and economic hardship over the course of its history. Most of the migrants were illegal migrants who were asylum seekers and refuges to escape from political instability, famine and persecution. The legal international labor migration was restricted during the military government of Ethiopia. However, this was changed when the FDRE government came to power since 1991. In addition, the FDRE constitution has allowed to the free movement of people.

Abdu and Girum (2010) witnessed the booming of labour migration from Ethiopia, particularly to the Middle East and RSA. Ethiopian labour migration to the Middle East reached its peak after Saudi Arabia began to massively recruit domestic workers from Ethiopia following its disagreements over wages and violations of the human rights of migrant laborers (domestic workers) with the governments of Indonesia, Sri Lanka and the Philippines. During the time, large numbers of Ethiopian women have been migrated to the Middle East in illegal way. As a result, they have been become victimized by human trafficking. Although, some have heard of

cases of abuses of migrant workers before they leave, most of them were not informed of the living condition awaiting them in the country of destination. Most of these women end up as modern day slaves.

Dejeneet al.(2011) examined that Poverty and lack of opportunities, failure in educational endeavors, gender stereotypes, and the "culture of migration" are critical factors behind migration and human trafficking. Most of the economic factors are related to low employment opportunities at the local level, low wage rates, low income, impoverished life and limited access to basic means of production such as land and credit facilities. Nowadays, many skilled and unskilled Ethiopians migrate to different countries legally and illegally looking for better economic opportunities.

Emerta et al. (2010) proposed that, Ethiopia lost a substantial number of skilled man power at different periods. "Ethiopia lost large numbers of graduates who have not returned after study abroad. In 2003, Ethiopians were the second largest group of immigrants to the US and they have been in the top four countries since at least 1990". Furthermore, from 1980-1991, it lost 74.6% skilled man power.

Migration has negative consequences in Hadiya zones. One of the negative consequences is that the zones have been losing their human resource power. Although, it has not created a significant professional scarcity in the zones currently, professionals mainly teachers have been leaving their jobs and migrating to the RSA. The negative consequence of teacher attrition cannot be denied that the vacuum created by emigrating teachers compromises the ability to provide quality education to future generations. Even though the problem was highly intensified in these grade levels, the other grade levels students also have been victimized by migration dream. Surprisingly, students' even children's dream is to migrate to the RSA like children commonly dream to be a doctor, an engineer and the like.

Tsedeke L. and Ayele T. (2017) tried to investigate socio-demographic and economic characteristics of a household on international migration and to estimate small area proportions at district and enumeration area level. A total of 2288data are collected from sixteen randomly sampled districts in Hadiya and Kembata-Tembaro zones in Southern Ethiopia. Several versions of the binary logistic mixed models, as special cases of the generalized linear mixed model, are analyzed and compared. The findings of the study reveal that about 39.4% of the households have at least one international migrant, and the rest 60.6% have no such migrants.

From analysis of the final model, it is found that the likelihood of a household of having international migrant increases with head's age and family size. An increase of family size by one person increases the log odds of having migrant by 0.131 indicating that large family size is one of the determinants for migration in the study area. The migration prevalence varies among the zones, the districts and the enumeration areas. Household characteristics: age, educational level and occupation of head, and family size are determinants of international migration. Community based intervention is needed so as to monitor and regulate the international migration for the benefits of the society

2.4Reviews on Statistical Model

Tsedeke Lambore and Ayele Taye (2017) investigate socio-demographic and economic characteristics of a household on international migration and estimate small area proportions at district and enumeration area level in Hadiya and Kembata-Tembaro zonal areas, Southern Ethiopia by using logistic mixed regression model. The findings of the study reveal that about 39.4% of the households have at least one international migrant, and the rest 60.6% have no such migrants.

Bereket et.al (2011) aimed to assess hazards and associated factors among returned migrants living at Bati Woreda, Amhara National Regional State and used cross-sectional survey with a sample size of 390 returnees was made in five kebeles of the woreda using a structured questionnaire. Multiple logistic regression analysis was used to assess the relative importance of associated factors. Socio-economic variables with respect to age, education and unemployment were predictors of the migration phenomenon. Migrants were highly exposed to health hazards because of the illegal routes of the movement. Socio-economic variables with respect to age, education and unemployment were predictors of the migration phenomenon. Providing safety protection from potential health hazards before leaving the country, proper social and psychological rehabilitation of returnees is recommended.

Nigatu and Ansha (2007) determined some socio-economic and demographic correlates play a pivotal role in the out-migration of masses from villages. The study was based on primary data collected from 1,258 households selected from four of the most populous districts (zone) of southern Ethiopia. The study has employed a multiple analysis to estimate the socio-economic correlates of out-migration in Southern Ethiopia. While the dependent variable of the study is

"whether at least one household member out-migrated during the past10 years or not", about eight important independent variables were regressed against the response variable. The rural out-migration in Southern Ethiopia is related to some non-economic and economic factors, among which lack of land, food insecurity, household size pressure, educational and religious status.

Clark G. and Valerie M. (2012) were focused on the possibility that climate change will displace large numbers of migrants in the developing world. He used a unique longitudinal data set from the Ethiopian highlands to investigate the effects of multiple measures of drought on the labor and marriage related mobility of men and women over a ten-year period. The result of multivariate studies indicate that men's labor migration increases with drought and those land-poor s households are most vulnerable. However, marriage-related moves by women decrease with drought, suggesting a hybrid narrative of environmentally induced migration that recognizes multiple dimensions of adaptation to environmental change.

In view of the reviewed literature, in this study the most important expected factors that influence migration of people were examined using count regression models. Based on reviewed literature as stated in this section, the explanatory variables which are expected to affect numbers of migrants per household are categorical and continuous.

CHAPTER THREE

3. DATA AND METHEDOLOGY

3.1 Description of the Study Area and Target Population

This study encompasses rural peasant association and rural-urban town of Soro and Misha district of Hadiya Zone from the southern parts of Ethiopia. Hadiya Zone is one of the major zones in Southern Nations, Nationalities and Peoples Region (SNNPR), located in the northern part of the Southern region of Ethiopia. The zone is divided into 11 Woredas, including Hossana town administration. These are Gombora, Lemo, Soro, East Badawacho, Misha, Shashogo, Duna, Gibe, West Badawacho and Anlemo. Soro District is one of the districts in Hadiya zone. Its capital is Gimbichu, which is 32 kilo-meters from Hosanna, capital of Hadiya zone and 201 kilo meters west of Hawassa town, capital of SNNPRS.

Concerning the information of climate, the Soro District has three agro-ecological zones sharing, Dega, Weyna dega, and Kolla. The annual average temperature of the Soro district is 20 - 30°Celsius. According to Soro District's finance and economic office report, it has49kebeles including two rural-urban kebeles with an estimated population of 214,630 peoples. Misha District is also one of the districts in Hadiya zone. Its capital is Morsito, which is 27 kilo-meters from Hosanna, capital of Hadiya zone. Misha woreda has 33 kebeles including two rural-urban kebeles Geja and Morsito.

Concerning the information of climate, the Misha District has three agro-ecological zones namely Dega, Weyna Dega, and Kolla. In both districts many of its residents are primarily ethnic group of Hadiya, followed by Kembata and Danta origins they all live together peacefully. While the majority of its residents speak Hadiyissa as their first language, some speak other language such as Amharic, Kizigna etc...The population in both districts is predominantly Protestant, with some Muslims, Orthodox and Catholics. The target population of this study is peoples of both sexes in Soro and Misha districts. They are contacted in households where there are people that have moved to the RSA out-migrants, return migrants from South Africa; and non-migrants who have no migration experience to South Africa. Information about out-migrants, who already moved to RSA, was gathered using proxy respondents mainly from their

families at home. Information about out-migrants who prepared themselves to migrate in short time with their networks or a smuggler was gathered from themselves.

3.2 Data Sources

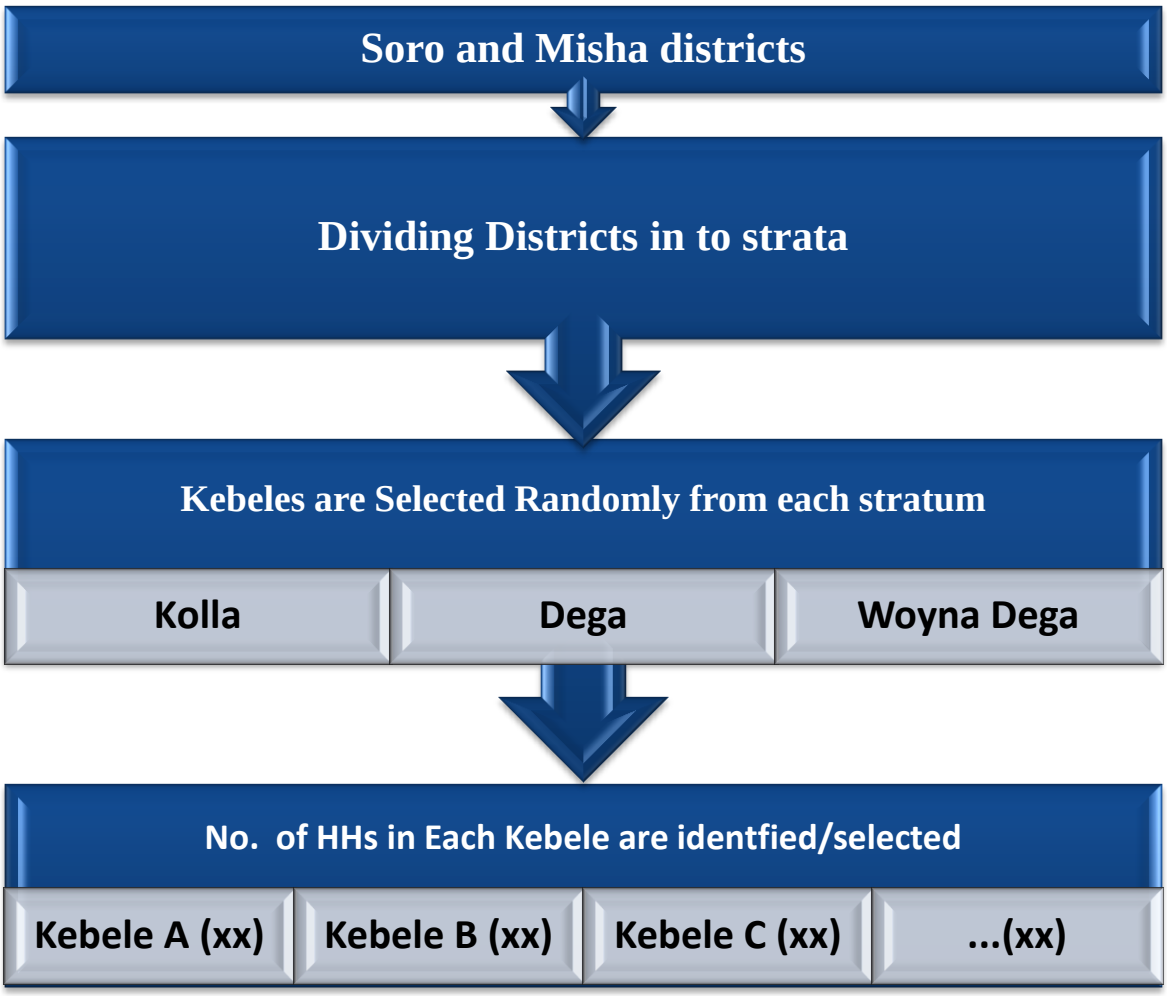
The research instruments that were employed under this study were mainly a structured questionnaire to selected households in Soro and Misha woreda/districts. The districts were selected based on key informant interviews and some relevant literatures on migration. Both qualitative and quantitative data pertaining to demographic; socio-economic; cause and consequence of migration, attitudinal and their problem on the journey of respondents (migrants)including out migrants, non-migrants and returnees.

Also Secondary data were used for additional information thoroughly whenever necessary. The main purpose of doing a qualitative data in this study is that most of the challenges and harsh experiences encountered by both out migrants and returnees are better addressed.

3.3 Sampling Technique

The sampling techniques which were used in this study were multi-stage random sampling techniques. Selections of eligible participants were done as follows: The districts were divided in number of groups (strata) which were used to select kebeles randomly at first stage to avoid the bias in the selection process. Secondly, the numbers of household were selected by using systematic random sampling method from each kebeles.

The following is the sampling design for this study:



Source: Self Compilation

Figure 3.2: Sampling Design

The sample size was distributed to the sampling units proportionate to the size of the households in the kebeles. Number of households in each Kebeles obtained from the Kebele office. Study households were selected from each Kebele through systematic sampling from a random start point. The sampling interval of households in each Kebele was determined by dividing the total number of households to the allocated sample size.

The first interviewed household was randomly selected by a lottery system from the Kebele house number register, using a number between one and the sampling interval. Where, sampling interval is the ratio of total number of households in the kebeles to allocated sample size to be selected from that kebele. The subsequent households to be included in the study were identified systematically through house-to-house visits, each time adding the sampling interval to the previous number.

The minimum sample size required from Misha district was 372 and from Soro district was 382. At last total sample of 754 households are to be selected from the study area.

3.4 Methods of Data collection

The research instruments that were employed under this study were questionnaire. Primary data were collected through structured questionnaire to selected households in districts. The questionnaire was designed to have both qualitative and quantitative data relevant to a slightly modified form of determining variables of migration, some demographic, socio-economic, push factors pull factors, environmental cause, positive consequence, negative consequence and harsh or problem on the way were constructed. The data were collected with the help of trained enumerators and the investigator. In order to minimize the errors in data collection that may be introduced by the enumerator training were provided to ensure that the questionnaire is understood by the enumerator. The questionnaire is found in Appendix

3.5 Variables of the study

The dependent and independent variables were selected based on experiences from the available similar studies.

3.5.1 The Response Variable

The response variable of this study is a count variable, the number of migrants from each household.

3.5.2 Explanatory Variables/Factors

Based on the reviewed literature, some of the common predictors that are expected to influence the numbers of migrants per households in the study area are given below.

Table3.1. Description of the predictor variables.

Variable name	***Variable Description and coding***
1. Age of household head(year)	<45(0) 45-60(1) >60(2)
2. Sex of household head	Male(1) and Female(0)
3. Average income of HH	Total average income of the household member(continuous)
Main economic activity or job of household head.	Merchant(0) Agriculture(1) Admn/clerical(2) Other(3)
4. The highest level of education attained by household head	Illiterate(0) Primary(1) Secondary(2) Higher(3)
Religion of household head	Protestant(0) Orthodox(1) Muslim(2) Other(3)
5. Previous Migration status of HHH Except to RSA	Migrated(1) Not migrated(0)
6. HH size	Number of household member(discrete)

7. HH land size	Household land size in hectare(continuous)
8. Illiteracy ratio	Ratio of illiterate household members ; number of illiterate member/number of literate(continuous)
9. Dependency ratio	Ratio of dependent household members to working age household members [number of children aged less than 15 year + number of elderly aged greater than65/number of working age adult](continuous)
10. Residence place of household	Rural(0) Urban(1)
11. Home ownership of household	Own(0) By rent(1) Other(2)

Methods of Data Analysis

3.6. Count Regression Model

Count data regression models have been widely used in statistics to model response variable a form of count. In this study, the variable of interest is a count variable. When the response or dependent variable (number of migrants per house in this study) is a count (which can take on non-negative integer values, it is appropriate to use non-linear models based on non-normal distribution to describe the relationship between the dependent variable and a set of predictor variables. For count data, the standard framework for explaining the relationship between the outcome variable and a set of explanatory variables includes the Poisson and Negative Binomial regression models. Unlike linear regression, count data regression models have counts as the response variable that can take only nonnegative integer values.

3.6.1 Poison Regression Model

Poisson regression model is commonly used to study the association between count outcome variable and covariates. It is the baseline model for count data analysis. However, its restrictive assumptions often make it inadequate in real-life applications. There are two strong assumptions

for Poisson model to be checked: one is the events that occur independently over time or exposure period, the other is the conditional mean and variance is equal. A Poisson regression model allows modeling the relationship between a Poisson distributed response variable and one or more explanatory variables. It is suitable for modeling the number of events that occur in a given time period or area. In practice, counts have greater variance than the mean described as over-dispersion. This indicates Poisson regression is not adequate. There are two common causes that can lead to over-dispersion: additional variation to the mean or heterogeneity, a negative binomial model is often used and other cause counts with excess zeros or zero-inflated counts, since the excess zeros will give smaller mean than the true value, this can be modeled by using zero-inflated Poisson (ZIP) or zero-inflated negative binomial (ZINB)(Cameron and Trivedi, 1998).

Poisson Regression Model provides a standard framework for the analysis of count data. Let Y_i represent counts of eventsoccurring in a given time or exposure periods or area with rateμ_i,Y_i are Poissonrandom variableswith probability mass function (pmf) given below**:**

$$P(Y_i=y_i, \mu) = \frac{e^{-\mu}\mu^{y_i}}{y_i!}\ y_i=0,1,2,3\ldots$$

μ>0, i= 1, 2 ...n where, Y denotes the number of migrants per household, $\mu$is the rate parameter which is non-negative and it is given as:

$E[y_i] = \mu_i = \exp(X^T\beta)$ where $X^T = (1, X_{1i}{}^T, X_{2i}{}^T \ldots)$

And β=p+1 dimensional column vector of unknown parameters to be estimated and p is number of predictors. The estimation is undertaken by using maximum likelihood method. The first two moments of the Poisson random variable Y are E[Y] = μ and V[Y] = μ.If both are equal that means conditional mean is equal to conditional variance this shows the well-known equ-dispersion (equal mean and variance) property of the Poisson distribution.

3.6.2Estimation of Parameter of Poisson Regression Model

The Poisson regression model is a nonlinear regression model. It is derived from Poisson distribution by allowing the rate parameter μ to depend on covariates. The most commonly used formulation is the log-linear regression model as given below:

Log (μ_i*)=*$X_i^T \beta_i$where, X^T= $(X_1, X_2, X_3 \ldots X_p)$ is vector of the explanatory variable andβvector of unknown regression coefficient.

The regression parameters are estimated using the maximum likelihood estimation. The likelihood function of the Poisson model based on a sample of n independent observations is given by:

$$L(\beta) = \prod_{i=1}^{n} \frac{e^{-\mu} \mu^{y_i}}{y_i!}$$

The log-likelihood function is:

$$l = \log(L(\beta)) = \sum_{j=1}^{n} [-\mu_i + y_i * \log(\mu_i) - \log(y_i!)]$$

The likelihood equation for estimating the parameter is obtained by taking the partial derivations of the log-likelihood function and setting them equal to zero. Thus, we obtain the first derivatives of function with respect to the underlying parameters as follows:

$$\frac{\partial l(\beta)}{\partial \beta_i} = \sum_{i=1}^{n} (y_i - \mu_i) x_{ij}$$

3.6.3 Limitations of Poisson regression model

Poisson regression model is appropriate for modeling count data when successive events occur independently and at the same rate. However, in practice, the features of data often violate these assumptions. Usually, the variance of count data exceeds its mean, resulting in over-dispersion. This may be due to unobserved heterogeneity, as the rate parameter is not only determined by a deterministic function of X^Tbut also by a random (unobserved) component. The over-dispersion may also result from excess zeros, which happens when observed zeros are much more than that predicted by the assumed distribution. Moreover, over-dispersion will result in deflated standard

errors of parameter estimates and therefore inflated t-statistics. Thus, it is always necessary to conduct a test of over dispersion after the development of Poisson regression.

If $E[y_i] < var\ [y_i]$,then we speak about over- dispersion, and when $E[y_i] > var[y_i]$we say we have under-dispersion. Next, we employed two tests of over dispersion where the Null Hypothesis (H_o) is mean and variance of the response variable are equal against the Alternative Hypothesis (H_1) variance exceeds the mean. There are two basic criteria commonly used to check the presence of over dispersion:

1. Deviance, D(y,$\hat{\mu}$), is given by:

$$D(y,\hat{\mu}) = -2\left[\sum_{i=1}^{n} y_i log(\frac{y_i}{\widehat{\mu_i}}) - y_i - \hat{\mu}\right]$$

Where, y is the number of events, n is the number of observations and$\hat{\mu}$, isthe fitted Poisson mean.

2. Pearson chi-square test, χ^2 is given by:

$$\chi^2 = \sum_{i=1}^{n} \frac{(y_i - \widehat{\mu_i})^2}{\widehat{\mu_i}}$$

Over-dispersion may be a result of higher occurrence of zero counts and subject heterogeneity. If the model fits the data, both deviance and Pearson Chi-square statistics divided by the degrees of freedom are approximately equal to one. Values greater than one indicate over dispersion, while values smaller than one indicate under-dispersion. It is possible to account for over-dispersion with respect to the Poisson model by introducing a scale (dispersion) parameter into the relationship between the variance and the mean (Pedan, 2001). Another way of checking the presence of over-dispersion is a statistical test of the hypothesis:

$$H_o: \alpha = 0 \ \ vs\ H_1 : \alpha > 0$$

If P-value of LRT α < (level of significance), then there is over-dispersion and the Negative Binomial model is preferred. The Negative Binomial Regression Model is more appropriate for over-dispersed data because it relaxes the constraints of equal mean and variance. In the general,

Poisson Regression Model, we think of μ_i as the expecteddesired number of migrants from i^{th}household and the total number migrants from i^{th} household is N_i. This means, parameter will depend on the population size and the total number of migrants from individual household. Thus, the distribution of Yi can be written as:

$$Y_I \sim \text{Poisson}\ (N_i, \mu_i)$$

Where N_i are the total migrant of i^{th} household and μ_i is the logarithm of the migrant is introduced in the regression model as an offset variable. By including:

$$Log\ (\mu_i) = Log\ (N_i) + X_i^T \beta$$

The link between the expectation of the dependent variable and the linear predicator is a logarithmic function and the linear predictor contains a known part or offset. This allows for estimation of maximum likelihood, standard errors and the likelihood ratio goodness of fit chi-square statistics (Agresti, A.2007). The model suggests that both set of the parameters are dependent on the covariates.

Since *Log* (N_i) is a constant, any variation in the coefficients of the independent variables will show up affecting the dependent variable and not the number of household. The procedure therefore allows us to obtain the maximum likelihood regression coefficients that can be easily interpreted in terms of differentials in the dependent variables. Using the Negative Binomial Regression procedure, several regression equations are estimated to the relationship between non migrant changes when control variables earlier mentioned are introduced. Results from the Negative Binomial Models are some times better expressed on more convenient scale.

3.6.4 Negative Binomial Regression Model

The NB Regression Model is used when count data are over dispersed (i.e when the variance exceeds the mean). Over dispersion, caused by heterogeneity or an excess number of zeros (or both) to some degree is inherent to most Poisson data. By introducing a random component into the conditional mean, the Negative Binomial Regression Model addresses the issue of over-dispersion. However, it equally models both zero and nonzero counts, which might result in a

poor fit for data with excessive number of zeros. Therefore, it is always necessary to check the proportion of zero counts before developing a Negative Binomial Regression Model.

This study used the likelihood ratio test to determine the more appropriate model between the Poisson Regression and Negative Binomial Regression. Hilbe(2011) used Negative Binomial Regression to Model over dispersed Poisson data. When the Negative Binomial is used to model over-dispersed Poisson count data, the distribution can be thought of as an extension to the Poisson Model. The Negative Binomial Regression Model uses a log link function between the dependent variable (number of migrants per household) and independent variables. The only difference between the Poisson and the NB lies in their variances, regression coefficients tend to be similar across the two models, but standard errors can be very different. The NB regression model is:

$$p(y_i, \mu_i, \alpha) = \frac{\Gamma(y + \alpha^{-1})}{y!\,\Gamma(\alpha^{-1})}\left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu}\right)\left(\frac{\mu}{\alpha^{-1} + \mu}\right)^{y}$$

$y_i > 0 \;\; and \; \alpha > 0$with mean and variance are given by:

$E(y_i) = \; \mu_i = exp^{(X_i{}^T \beta)} and \; var(y_i) \; = \mu_i(1 + \alpha\mu_i)$

Where, $\boldsymbol{\alpha}$shows the level of over-dispersion and Γ (.) is the gamma function. If α =0 NB Regression Model will reduce to Poisson Regression Model. Oftendata will show over-dispersion (Variance > mean) or under-dispersion (Variance< mean). With over-dispersed data we may well use the Negative BinomialRegressionmodel.This Model adds unobserved heterogeneity by specifying

$E(y_i) = \; \mu_i = exp^{(X_i{}^T \beta)}$where$X_i{}^T$is 1xp row vector of covariate (including an intercepts), p is the number of covariate in the model and px1 column vector of unknown regression parameters.The likelihood function of the NB model based on a sample of n independent observations is given by:

$$L(y_i, \mu_i, \alpha) = \prod_{i=1}^{n} \frac{\Gamma(y + \alpha^{-1})}{y!\,\Gamma(\alpha^{-1})}\left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu}\right)\left(\frac{\mu}{\alpha^{-1} + \mu}\right)^{y}$$

The log-likelihood function ℓ of NB regression model is:

$$l = \sum_{i=1}^{n} \{-\log(y_i!) + \sum_{k=1}^{y_i} (\alpha y_i - \alpha k + 1) - \left(y_i + {}^{1}/_{\alpha}\right) \log(1 + \alpha\mu_i) - y_i \log(\mu_i) \Bigg\}$$

Where, $\frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{y_{i!}\,\Gamma\left(\frac{1}{\alpha}\right)} = \prod_{k=1}^{y_i} \left(y_i + \frac{1}{\alpha} - k\right) = \alpha^{-y_i} \prod_{k=1}^{y_i} (\alpha y_i - \alpha k + 1\,)$

For estimating regression coefficients and dispersion parameter the Newton-Raphson iteration procedure is applied like in the Poisson model.

3.6.5 Zero-inflated Regression Model

Real-life count data are frequently characterized by over-dispersion and excess zeros. Zero inflated count models provide a parsimonious yet powerful way to model this type of situation. Such models assume that the data are a mixture of two separate data generation processes: one generates only zeros, and the other is either a Poisson or negative binomial data-generating process Count data that have an incidence of zeros greater than expected for the underlying probability distribution can be modeled with a zero inflated distribution. The population is considered to consist of two subpopulations. Observations drawn from the first subpopulation are realizations of a random variable that typically has either a Poisson or Negative Binomial distribution, which might contain zeroes.

3.6.6 Zero- inflated Poisson Regression Model

Suppose the mean of the underlying Poisson distribution is μand the probability of an observation being drawn from the constant distribution that always generates zeros is ω_i. The parameterω_i is often called the zero inflationprobability (Agresti, A., 2002).

The probability distribution of a zero inflated Poisson random variable Y is given by:

$$P(Y_i = y_i) \begin{cases} \omega_i + (1 - \omega_i)e^{-\mu}, & y_i = 0 \\ (1 - \omega_i)\dfrac{e^{-\mu_i}\mu_i^{y_i}}{y_i!} & y_i = 1, 2, \ldots \quad 0 \leq \omega_i \leq 1 \end{cases}$$

The mean and variance of Zero-inflated (ZIP) distribution is given as

$E(Y_i) = (1 - \omega_i)\mu_i$ and

$\mathrm{Var}(Y_i) = E(Y_i)((1 - \omega_i)\mu_i)$.

The excess zeros are a form of over-dispersion and fitting a zero inflated Poisson model can account for the excess zeros, but there are also other sources of over-dispersion that must be

considered. If there are sources of over-dispersion that cannot be attributed to the excess zeros, failure to account for them constitutes a model misspecification, which results in biased standard errors. In ZIP Models, the underlying Poisson distribution for the first subpopulation is assumed to have a variance that is equal to the distribution's mean. If this is an invalid assumption, the data exhibit over-dispersion (or under-dispersion). (Agresti, A., 2002) A useful diagnostic tool that can aid in detecting over dispersion is the Pearson chi-square statistic defined as

$$\chi^2 = \sum_{i=0}^{n} \frac{(y_i - \mu_i)^2}{V(\mu_i)}$$

Comparing the computed Pearson chi-square statistic to an appropriate chi-squared distribution with n-p df constitutes a test of over-dispersion. If over dispersion is detected, the ZINB Model often provides an adequate alternative

3.6.7 Zero-Inflated Negative Binomial Regression Model

Zero-inflated Negative Binomial Regression is for modeling count variables with excessive zeros and it is usually for over-dispersed count outcome variables. Furthermore, theory suggests that the excess zeros are generated by a separate process from the count values and that the excess zeros can be modeled independently. The probability distribution of a Zero inflated Negative Binomial random variable Y is given by:

$$P(Y_i = y_i) = \begin{cases} \omega_i + (1 - \omega_i)(1 + \alpha\mu_i)^{-1/\alpha} & , y_i = 0 \\ (1 - \omega_i)\dfrac{\Gamma(y_i + \frac{1}{\alpha})}{y_i!\,\Gamma(\frac{1}{\alpha})}(1 + \alpha\mu_i)^{-1/\alpha}(1 + \dfrac{1}{\alpha\mu_i})^{-y_i} & , \quad y_i > 0 \end{cases}$$

Where $\alpha > 0$ is a dispersion parameter and is assumed not to depend on covariates and $\mu_i$is the mean of the underlying negative binomial distribution. The mean and variance of the ZINB model are given by:

$$E(Y_i) = (1 - \omega_i)\mu_i \text{ and}$$
$$Var(Y_i) = (1 - \omega_i)(1 + \omega_i\mu_i + \alpha\mu_i)\mu_i$$

The ZINB distribution is not a standard GLM type exponential family distribution, even when the over dispersion parameter is known, and standard GLM fitting methods are not applied. To

obtain the parameter estimates of ZINB regression models$\widehat{\alpha}, \widehat{\beta}, \ \hat{\gamma}$, the Newton-Raphson method can be used.

The log likelihood function$l = l(\alpha, \mu_i, w_{i\,;}\, y)$, for the ZINB Model is given below.

$$\ell = \ell(\alpha, \mu_i, w_{i\,;}\, y)$$

$$= \sum_{i=0}^{n} \left\{ I_{(Y_i=0)} \log\left(w_i + (1-w_i)(1+\alpha\mu_i)^{\frac{-1}{\alpha}}\right) \right.$$

$$\left. + I_{(Y_i>0)} \log[(1-w_i) \frac{\Gamma\left(Y_i + \frac{1}{\alpha}\right)(1+\alpha\mu_i)^{\frac{-1}{\alpha}} \left(1+\frac{1}{\alpha\mu_i}\right)\left(1+\frac{1}{\alpha\mu_i}\right)^{-yi} \ ]}{Y_i!\,\Gamma\left(\frac{1}{\alpha}\right)} \right\}$$

Since, $\frac{\Gamma\left(Y_i+\frac{1}{\alpha}\right)}{Y_i!\Gamma\left(\frac{1}{\alpha}\right)} = \prod_{k=1}^{Y_i}\left(Y_i + \frac{1}{\alpha} - k\right) = \alpha^{-Y_i}\prod_{k=1}^{Y_i}(\alpha Y_i - \alpha k + 1)$

Newton-Raphson iteration procedure can be used for estimating the parameter of ZINB Regression Models.

3.7 Goodness of fit tests

3.7.1 Likelihood Ratio test

The Likelihood ratio test is a test of a null hypothesis against an alternative based on the ratio of two log-likelihood functions. The likelihood ratio test is a test of the overall model. The overall test statistic for likelihood ratio test is given as:

Likelihood ratio test: $G^2 = -2(l_{null} - l_k) \sim x^2_{p-1}$

This statistic is called the likelihood-ratio test statistic.

Where:l_{null} is the log-likelihood of the null model and

$l_k$is the log-likelihood of the full model

Comprising k predictor, p is number of parameters and $x^2_{p-1}$is a chi-squaredistribution with p-1 degree of freedom. If the test statistics exceeds the criticalvalue, the null hypothesis is rejected. That means the overall model is significant. In this study, to compare Poisson and NB regression models and also ZIP with ZINB regression models, we used significance of dispersion parameter and likelihood ratio (LR) test as criterions. The statistic of likelihood ratio test is given by the following equation:

LRT_α=-2(LL_1-LL_2)

This statistic has a Chi-squared distribution with 1 degrees of freedom and LLis log-likelihood. If the statistic is greater than the critical value then, the model2 is better than the model 1.

3.7.2 Vuong Test

The Vuong test is a non-nested test that is based on a comparison of the predicted probabilities of two models that do not nest (Vuong, 1989). That means Vuong test statistics are needed to provide the appropriateness of zero-inflated models against the standard count models. For instance, comparisons between zero-inflated count models with ordinary Poisson, or zero-inflated negative binomial against ordinary negative binomial model can be done using Vuong test. This test is used for model comparison. For testing the relevance of using zero inflated models versus Poisson and NB regression models, the Vuong statistic is used.

The formula for these is given below:

$$m_i = \log[\, p_1\left(\frac{y_i}{x_i}\right)] - \log[\, p_2\left(\frac{y_i}{x_i}\right)]$$

Where, $p_1\left(\frac{y_i}{x_i}\right)$ and $p_2\left(\frac{y_i}{x_i}\right)$ are probability mass functions of zero-inflated and Poissonor NB models respectively. In general $p_N(\frac{y_i}{x_i})$ is the predictedprobability of observed count for case i from model N, then the Vuong teststatistic is simply the average log-likelihood ratio suitably normalized.The test statistic is:

$$V = \sqrt{n}\frac{\frac{\sum_{i=1}^{n} m_i}{n}}{\sqrt{\frac{\sum_{i=1}^{n}(m_i-\bar{m})^2}{n}}} = \frac{\sqrt{n}}{s_m}\bar{m}$$

Where, $\bar{m}, n \; and \; s_m$, are mean, sample size and standard deviations respectively.

The hypotheses of the Vuong test are:

$$H_0: E[m_i] = 0 H_1: E[m_i] \neq 0$$

The null hypothesis of the test is that the two models are equivalent. Vuong showed that asymptotically, V has a standard normal distribution. As Vuong notes, the test is directional (vuong, 1989)

- If V >$Z_{\alpha/2}$, the first model is preferred.
- If V < -$Z_{\alpha/2}$, the second model is preferred.
- If $|V|$<$Z_{\alpha/2}$, none of the models are preferred

3.7.3 AIC and BIC

AIC and BIC are goodness of fit criteria used for model selection. The likelihood ratio test was used to compare the Poisson model and NB model. Many Monte-Carlo simulations indicate that the BIC and AIC selection criteria need to be used together [Dalrymple et al (2003) and Wang et al (1996)]. The model with smallest value of AIC or of BIC is preferable. Selecting an appropriate model is often based on a standard likelihood information criteria, for example,

Akaike information criteria (Akaike, 1973) or Bayesians information criteria abbreviated by AIC and BIC, respectively, where:

AIC=-2*log likelihood +2k

BIC=-2*log likelihood + k $\ln(n)$ Where, k = number of parameters and n = number of observations.

3.7.4 Test for individual predictors

Let β denote an arbitrary parameter. Consider a significance test ofH_0: β = 0. The simplest test statistic uses the large-sample normality of theML estimatorβ. Let $SE(\beta)$ denote the standard error of β , evaluated by substituting the ML estimate for the unknown parameter in the expression for the true standard error. WhenH_0 is true, the test statistics

$Z = \frac{\hat{\beta}-\beta_0}{SE(\hat{\beta})}$, has approximately a standard normal distribution.

Equivalently Z^2 has approximately chi-squared distribution with df = 1.This type of statistic, which uses the standard error evaluated at the ML estimate, is called a Wald statistic. The Wald statistic is: $Z^2 = (\frac{\hat{\beta}-\beta_0}{SE(\hat{\beta})}^2)$ under H_0 is true Z^2 havechi-squared distributions with $df =$ 1 Likelihood-ratio tests are generally consideredbeing superior (Agresti, 2007)

CHAPTER FOUR

4. Results and Discussion

This chapter discusses results of the study showing how selected socio economic, and demographic factors/variables affect number of migrants to RSA among households in Soro and Misha districts. The statistical analyses were performed using South Texas Art Therapy Association (STATA) statistical package version 13, Statistical Package for Social Science(SPSS) version 23, and Microsoft-Excel.

4.1 variable selection

In this study we have used Stepwise variable selection which is a combination of backward elimination and forward selection to identify the predictors in the model. This was done on Poisson regression model as it is the bench mark for other count regression models. Stepwise selection method addresses where variables were added or removed with respect to the p-value in the process. Based on stepwise variable selection procedure; age of HHH, sex of HHH, Average income of HH, job of HHH, education level of HHH, religion, Migration status, households size, land size, illiteracy ratio, dependency ratio place of residence and home ownership were included in the models at 5% level of significant. Further analyses are made only on those variables which are found significant factors for the number of migrants per household of Soro and Misha districts in SNNPR, Ethiopia.

4.2 Descriptive Statistics

In order to have an overall picture of the distribution of the number of migrant in the house hold to RSA, all descriptive analysis were performed.

Table 4.1 Percentage of migrants and non migrants in some category of variables

	category	*Migration status*					
		non Migrant		*Migrant*			
		n_1	*%*	n_2	*%*	*Total*	*%*
District	*Misha*	*164*	*22.5%*	*208*	*26.6%*	*372*	49%
	Soro	*222*	*28.5%*	*160*	*22.4%*	*382*	51%
HHH religion	*protestant*	*246*	*33.5%*	*258*	*33.5%*	*504*	67%
	orthodox	*93*	*11.5%*	*88*	*12.5%*	*181*	24%
	Muslim	*37*	*5%*	*11*	*1%*	*48*	6%
	other	*10*	*1%*	*11*	*1%*	*21*	3%
home ownership	*own*	*239*	*32%*	*246*	*33%*	*485*	64%
	by rent	*117*	*16%*	*90*	*12%*	*207*	27%
	other	*30*	*4%*	*32*	*4%*	*62*	8%
Education level HHH	*illiterate*	*84*	*11%*	*78*	*10%*	*162*	21%
	primary	*131*	*22%*	*131*	*12%*	*262*	35%
	secondary	*111*	*15%*	*109*	*14%*	*220*	29%
	higher	*60*	*8%*	*50*	*7%*	*110*	15%
Ethnicity	*Hadiya*	*241*	*32%*	*261*	*35%*	*502*	67%
	Kembata	*66*	*9%*	*34*	*5%*	*100*	13%
	Silte	*30*	*4%*	*25*	*2.5%*	*55*	6.5%
	other	*49*	*6%*	*48*	*6.5%*	*97*	13.5%
place of residence	*rural* *urban*	291 95	*38.6%* *12.6%*	*270* *98*	*35.8%* *13%*	*561* *193*	74.4% 25.6%

The data on Table 4.1 show that26.6% of the households in Misha district and 22.4%of the households in Soro district have at least one migrant.

It is instrumental to see the role of ethnicity on migration since the current migration is dominated by one ethnic group—*Hadiya*. In the study area, this ethnic group accounts the highest share among all. Over67% of the respondent households belong to *Hadiya* ethnic group followed by *Kembata*(13.5%). The rest are *Siltie*(6.5%) and belonging to other ethnic groups is 13.5% (each sharing less than15%).Table4.1 shows the majority of the household heads attended primary education (35%) and 21% of them are illiterates. Among the respondent households the education status of HHH indicates that 29% of them attended secondary education and only 15%

of them have followed higher education. The highest percentage of primary education is found among non-migrant household (27%) compared to 12% among migrant households.

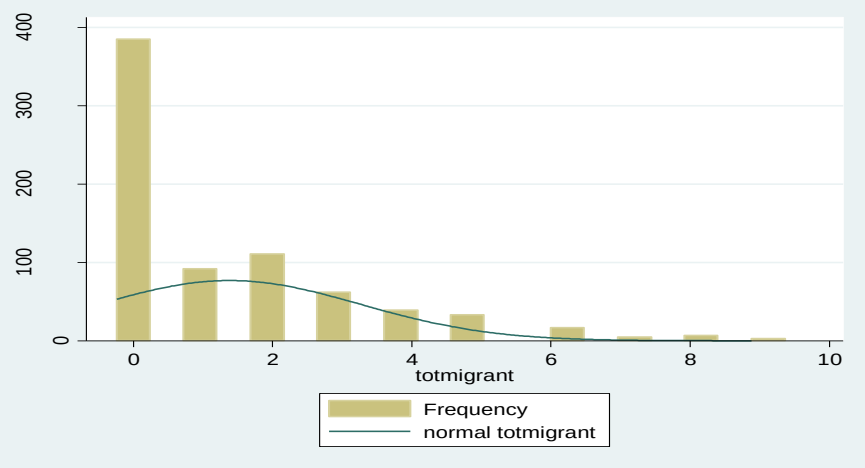
Table 4.1 shows Large percentages of respondent" household heads (67%) religion is Protestant followed by 24% Orthodox, 6% Muslim, and 3%others. Also 67% household leaves in their own house, 28% of households leave by rent and 8% of total household live neither by rent nor in their own house. Additionally, 74.4% of the household were living in rural whereas 25.6% of the household were living in urban.

Table4.2. Summary statistics of dependent variable

Variable	Category	observations	Mean	standard deviation
No. of Migrants				
	Soro	382	1.242	1.640
	Misha	372	1.338	1.747
	total	754	1.292	1.695

Table 4.2 shows the descriptive statistics of the dependent variable. The means number migrants household in districts Soro and Misha are 1.242 and 1.338, respectively, whereas the standard deviation of total migrants of household in the districts Soro and Misha are 1.64 and 1.747 respectively. The variance of the outcome variable (total number of migrant) for entire data, 2.84 is greater than mean 1.29, and the ratio $\frac{2.8561}{1.2923}$=2.223>1, the results suggesting over dispersion.

Figure 4.1 The below figure is histogram of number of migrants in study area.



Histogram totmigrant, frequency normal
(bin=41, start=0, width=0.4878)

As shown in Figure 4.1, the distribution of the number of migrants has a somewhat slowly decreasing tail and is skewed to the right.

Summary Statistics of Selected Variables

Summary statistics for the number of migrants in the household are presented in Table 4.2. The average number of migrants for the entire data is 1.29. The data revealed substantial variation in category of variables when controls are introduced

Table 4.3.Mean and SD of migrants among households

variables	***category***	***Mean***	***Std. Dev.***
tot-migrant		1.290	1.695
Sex HHH	*female*	0.931	1.477
	male	1.424	1.752
HHH religion	*Protestant*	1.415	1.789
	orthodox	1.039	1.271
	Muslim	0.750	1.839
	others	0.784	1.821
homeownership	*own*	1.268	1.608
	by rent	1.000	1.386
	other	2.435	2.615

Place of residence	*rural*	1.233	1.627
	urban	1.455	1.873
Ethnicity	*Hadiya*	1.363	1.675
	Kembata	0.750	1.234
	silt	1.164	1.500
	other	2.546	2.156
HHH migration experience	*non migrated*	1.000	1.389
	migrated	3.920	1.944
HHH education level	*illiterate*	1.333	1.694
	primary	1.275	1.607
	secondary	1.223	1.726
	higher	2.100	1.848
HHH job occupation	*merchant*	1.232	1.665
	agriculture	1.466	1.549
	Admn/clerical	1.406	2.015
	other/student	0.563	0.914

Sex of HHH: There was a difference in the average number of migrants by sex of household head. The mean of number of migrants was higher in male headed households (1.424) than in female headed households (0.96).

HHH religion: The average number of migrants was highest (1.41) among protestant household heads and lowest (0.75) among Muslim household heads. The average number of migrants was 1.04 among orthodox household heads whereas the average number of migrants was 0.784 for household heads who follow another religion.

Homeownership: There was a difference in the average number of migrants by home ownership. The mean of number of migrants was lowest (1) in households living in a rented house and highest (2.43) in households living neither in rented house nor their own house. Also the average number of migrants in the households owning their house was 1.26.

Place of residence: There was a difference in the average number of migrants by place of residence. The mean of number of migrants was higher in urban areas (1.45) than in rural areas

(1.23). The number of migrants among rural household is almost equal to among urban household.

Ethnicity: There was variation in the average number of the migrants by ethnicity. The average number of migrants was highest in other type ethnic group followed by *Hadiya* (1.36), *Silte* (1.16) and *Kembata* (0.75).

HHH education level: There difference in the number of migrants has been observed between household heads with no education and those with education. Household head with some education exhibited higher average number of migrants than those without education. That is, difference in the average number of migrants was observed between household head that have never been to school and those with secondary and higher education. Average desired number of migrants in household with no education was 1.33. Conversely, the average number of migrants of household with primary education, secondary education and higher education was 1.275, 1.223and 2.1respectively.

Table 4A1which found in appendix shows large percentages of respondent (86%)in migrant household claimed that migration changed their economy, 10% of respondent in migrant household claimed that there is no change in economy and over 4% of respondents in migrant household are not sure about for the change of their economy. Among entire respondents 86% of them cited that there was a positive consequence in their environment. Remittance and reduction in poverty shares large percentage of respondents claim. As respondents belief percentage of response for job creation and Diaspora benefit in the society were too small as compared to remittance and reduction in poverty.

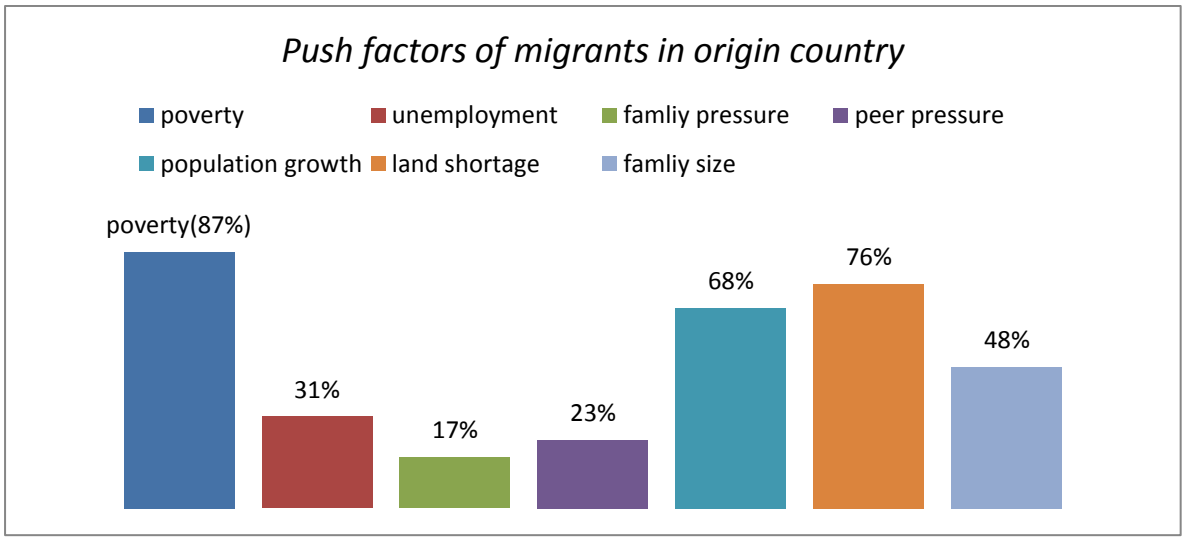
As mentioned in Table4A1 which found in appendix, among negative consequences, over 62%respondents claimed that death as a major problem in the journey. Similarly, among all respondents 49%, 71%, and 34% of the respondents claimed that lack of residence, lack of food and water and robbery respectively were the major factors in the journey. Based on the above table about 83% of the respondents cited that the peoples were migrating in irregular way and 17% of them claimed the way of migration was regular.

4.3 Push and pull factors of migration

4.3.1 Push factors

In addition to the result discussed above, respondents were asked on the main reasons for the migration of the peoples to South Africa and this is presented in Figure4.2. Over 87% of the respondent household claimed that their main reason they left their homeland was poverty. And 76% of respondent household cited land shortage, 68% claimed their reason is population growth, 48% claimed that their reason was large family size, 31% of respondents claimed that the main reason was unemployment,23% of respondent claimed that their reason was peer pressure and 17% of the all respondent cited their main reason to move where family pressure.

Figure4.2 Push factors of migrants in Origin country



Source: field survey (March- April 2018)

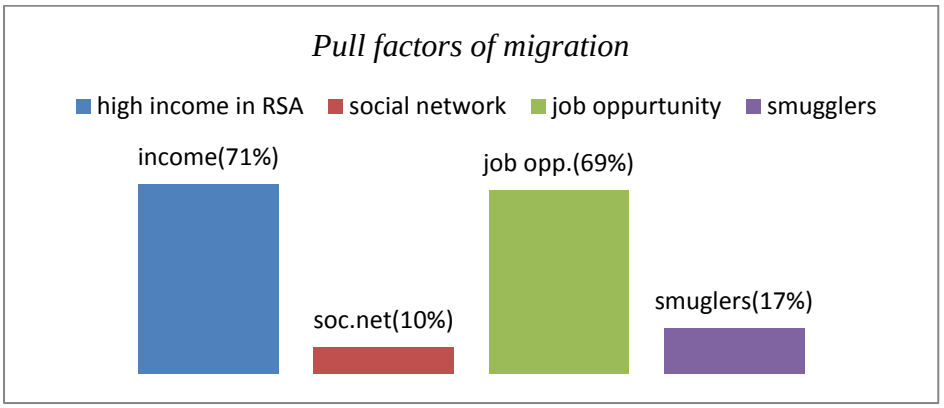
As mentioned above, among the push factors of migration, poverty has contributed the highest percentage for the migration in the study area. Another push factor that makes the area's people to migrate is unemployment. This is related to lack of various job opportunities other than practicing with small enterprises. The peoples have shortage of skills to create jobs in alternative livelihoods due to lack of the vocational training institutions in a district

4.3.2 Pull factors of migration

As presented in figure 4.3 below causes of migration among pull factors for people migration to RSA, high income in RSA has contributed to the highest percentage of about 71% for the

respondents and social network has contributed to the lowest percentage of about 10%.Over 69% of the respondents claimed that their main reason they left their homeland are of perceived better opportunities found in South Africa and 17% claimed that their main reason was smugglers.

Figure4.3: Pull factors of migrants



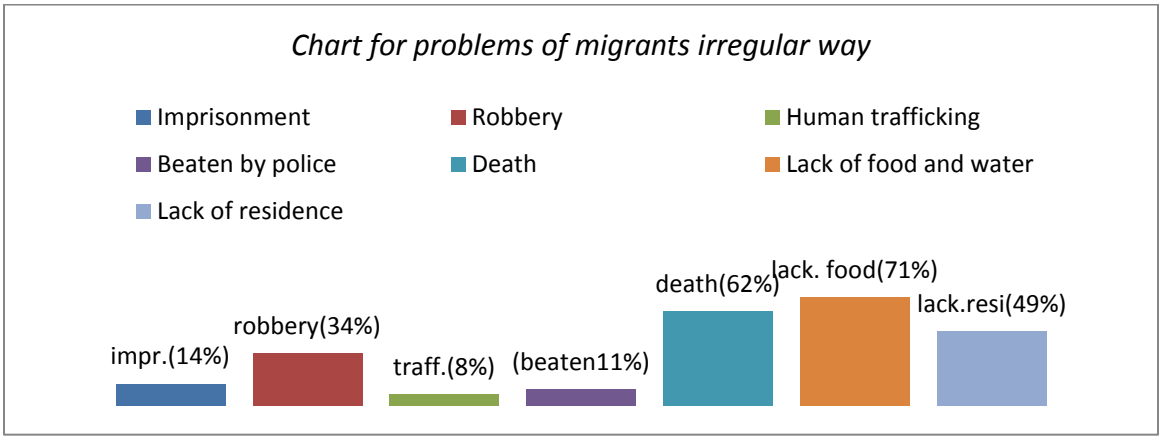
Source: field survey (March- April 2018)

The study found that the movement of peoples from southern Ethiopia to RSA is facilitated by a network of human smugglers and they work in association with several smugglers from abroad country.

4.3.3 Problems of Migration

Over 24% of respondents gave the response that the migration has no negative consequence on the journey to South Africa and about 74% of the respondents gave the response that the migration has negative consequence on the journey to South Africa like falling in to prison by lack of legal or documented transportation. From those respondent households 14% claimed that migrants were felt down to prison without any other criminal habits other than illegal movement. Next to imprisonment migration has danger or harsh impacts on migrant's life. About 62% of the respondents argued that migration has negative impacts on migrant's life (death) on the way. The study also reveals that 71% of respondent households claimed that migrants were encountered with problem of lack of food and water on the way (journey) to Republic of South Africa by different harmful conditions.

Figure4.3: Problems of Migrations



Source: field survey (March- April 2018)

About 49% of the respondents said that their journeys were harsh with unexpected negative consequences like lack of residences place. 11% percent of the respondent household reported that migrants were beaten by police and 8% of the respondent household claimed that migrants were encountered with problem of the human trafficking.

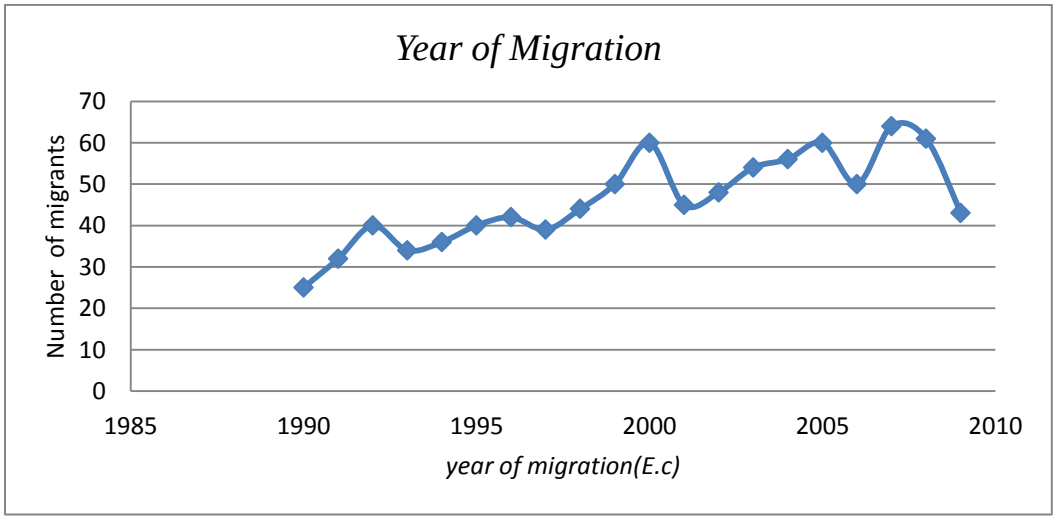
Based on TableA1, about 12% of respondents gave the response that the migration has no positive consequence on the journey to South Africa and about 86% of the respondents gave the response that the migration has positive consequence on the journey to South Africa. The entire respondent household also intended to identify the positive consequence on the environment by different positive results like flow of remittance, job creation opportunities, Diaspora benefits, decrement in poverty, improvement of social services and others. From the table4.4result, high percentage of the respondents household cited as a positive consequence were decrement in poverty(98%) followed by flow of remittance(95%), job creation opportunities(6%), Diaspora benefits(15%), Improvement of social services(28%)and other(3%) type of positive consequences.

As we seen above on positive consequence of migration there is also negative consequence on environment like Inequality among the people, Dependency on remittance, Shortage of labour force, Lack of job creation interest and Loss of migrant's life. From the Table4.4 result, high

percentage of the respondent households claimed that a negative consequence were loss of migrant's life(98%) followed by economic inequality among the people(64%),shortage of labor force (47%),Dependency on remittance(34%), lack of job creation opportunities(28%), and other(13%) type of negative consequences.

The negative consequences of migration is not bounded only above listed impacts but also mainly seen in Education or on students as presented in Table4.4.From the total selected respondent households 60% of respondents justify that migration has negative impacts on student's education. High percentage of the respondent households claimed that a negative consequences on students education were absenteeism (81%) followed by school dropout (69%), low achievement (28%) and brain drain (5%). Next to school dropout migration has negative consequence on education by following skilled or professional migrants that is "brain drain" which make next generation to be follow their former educated man power way and to reduce their education by choosing migration as their future vision. And the other impacts of migration on education are low achievement, Absenteeism and other related impacts of migration.

Figure 4.3 Number of migrants over the year.



Source: field survey (March- April 2018)

It can easily be observed from the figure that the number of migrants is increasing overtime. In 1990E.C, for example, the amount of migrants was 25 and this has increased significantly over the consequent years. The number of migrant was 60 in 2000 E.C, but increased with

fluctuations, reaching a maximum of 64 in 2007E.C year. This was the trend of number of migrants over the years.

Table4.4. Some characteristics of migrants

Variables	***Category***	***Obv.***	***Proportion***
		961	
Sex Migrants	Female	270	0.280
	Male	691	0.719
Age of Migrant	<15	39	0.040
	15-30	441	0.460
	31-45	337	0.350
	>45	144	0.149
Education Level	illiterate	48	0.049
	primary	163	0.169
	secondary	682	0.709
	higher	68	0.069
job occupation	agriculture	164	0.170
	Merchants	403	0.419
	Admn./clerical	134	0.139
	other	260	0.270
Migration experience	migrated	88	0.091
	not migrated	873	0.909

Source: field survey (March- April 2018)

The data on Table 4.5 show that 71.9% of the respondents are males and only 28.1% of them are females. The distribution of migrants among the sex shows slight variation. Looking at the migrants' age is necessary in understanding at what age migrants become more vulnerable to migration. The data on Table 4.5 shows that in all of the migrants aged 15-45 year take up the largest share. The very young and the very old adults found to be less likely to migrate. Generally, the migrants whose ageless than 15 years and the migrant age greater than 45 year are less likely to migrate while propensity of irregular migration is higher among age group 15-45 years.

Table4.5 shows the majority of the migrants are literates. The education status of migrants indicates that 70.9% of them attended secondary education, 16.9% of them attended primary education,4.9% of them were illiterate and only 6.09% of them have college diploma or above.

Also Table 4.5 shows the job occupational distribution of migrants where the largest percentages were engaged in trade (Merchants) (about42% of the total). About 17% of migrants were engaged in agriculture, 13.9% of them engaged in clerical and 27% of migrants engaged in other type of activity.

4.4. Count Regression Model Results

4.4.1. Parameter Estimation

Table4.6 presents the fitted models (with estimated coefficients and their corresponding standard errors) for the four count regression models considered (Poison, Negative Binomial, Zero-inflated Poison, and Zero-inflated Negative binomial models). Also presented in this table are over dispersion parameters alpha and its natural Logarithm. The results show that the over-dispersion parameter alpha is significantly different from zero indicating over-dispersion of the data. Moreover, the ratio of the Deviance and Pearson Chi-square statistic to their corresponding degrees of freedom are greater than one, indicating over-dispersion in the data and the Negative Binomial (NB) regression model is preferred over the Poisson model.

Table 4.5.: The results of over-dispersion test after fitting a Poisson regression

Statistics	**Value Deg. of freedom**	**Value/Deg. freedom**	**p-value**
Deviance test statistics	1219.434729	1.670	0.000
Pearson Chi-square statistic	1302.256 729	1.796	0.000

The fact that the Negative Binomial regression model is favored over the Poisson regression model has also been confirmed by the likelihood-ratio test. Since the value of this statistic is 36.24 with p-value less than 0.05, we reject the null hypothesis that there is equi-dispersion, and conclude that there is significant over-dispersion in the data.

When the assumption of the equality of variance and mean in the Poisson regression model is violated, over-dispersion occurs and the standard error estimates will be biased which leads to incorrect value of the test statistic. Since both AIC and BIC values were lower for the Negative Binomial model than for the Poisson model, the Negative Binomial model is preferred. The Negative binomial model accounts for the over-dispersion in the data.

Table4.6. Estimated Coefficients and Standard Error of the Poisson, NB, ZIP and ZINB models

Models	Poisson		NB		ZIP		ZINB	
Variables	Coef.	Robust Std. Err.	Coef.	Robust Std. Err.	Coef.	Robust Std. Err.	Coef.	Robust Std. Err.
Sex HHH								
male	0.323	0.113	0.327	0.115	0.323	0.06	0.3436	0.0864
Age HHH								
45-60year	0.077	0.174	0.081	0.269	0.072	0.069	0.071	0.0125
>60year	0.054	0.138	0.050	1.135	0.054	0.160	0.050	0.2650
HHH religion								
Orthodox	-0.062	0.114	-0.063	0.116	-0.068	0.125	-0.063	0.0902
Muslim	-0.387	0.289	-0.600	0.270	-0.367	0.104	-0.3496	0.2944
other	1.690	0.336	1.652	0.366	1.691	0.169	1.6795	0.3789
HH size	0.100	0.018	0.102	0.020	0.106	0.011	0.1102	0.0178
HH homeownership								
by rent	0.196	0.136	0.195	0.139	0.398	0.072	0.3932	0.1016
other	0.734	0.131	0.735	0.158	0.738	0.085	0.6915	0.1384
Ethnicity								
Kembata	-0.029	0.019	-0.037	0.205	-0.029	0.136	-0.0292	0.1863
silt	-0.681	0.194	-0.766	0.223	-0.825	0.105	-0.7903	0.2085
other	0.104	0.137	0.103	0.143	0.105	0.094	0.1055	0.1161
place of res								
urban	-0.217	0.092	-0.222	0.103	-0.218	0.059	-0.2511	0.0774
Job.HHH								
agriculture	-0.067	0.119	-0.065	0.113	-0.069	0.07	-0.6248	0.0856
Admn/clerical	-0.495	0.126	-0.498	0.141	-0.510	0.082	-0.5647	0.1032
other/student	-1.661	0.503	-1.617	0.493	-1.523	0.218	-1.662	0.4989
Land size HHH	-0.505	0.032	-0.505	0.038	-0.503	0.026	-0.5001	0.0358
Mig.exp.HHH								
non migrated	-0.363	0.098	-0.326	0.105	-0.353	0.061	-0.359	0.0874
Education level HHH								
primary	-0.046	0.122	-0.049	0.148	-0.045	0.089	-0.0491	0.1202
secondary	-0.047	0.119	-0.047	0.138	-0.049	0.082	-0.0427	0.1214
higher	-0.092	0.097	-0.930	0.212	-0.091	0.103	-0.0914	0.1811
Il.ratio	0.034	0.071	0.358	0.093	0.330	0.036	0.3424	0.0687
Dep. ratio	0.017	0.084	0.016	0.082	0.014	0.013	0.0222	0.0549
-cons	-1.953	0.058	-2.453	0.003	-0.102	0.141	-0.6186	0.0265
ln(hhsize)	1.000	*(exposure)*	1.000	*(exposure)*	1.000	*(exposure)*	1.000	*(exposure)*
/lnalpha			-0.855	0.220			-7.6000	0.6250
alpha			0.607	0.117			0.0005	0.0002

Note: *alpha is dispersion parameter. P-value<0.05 is statistically significant.*

4.4.2 Comparison of Models

How to choose the appropriate models for a specific study is a critical question in data analysis. Several criteria can be used to compare and select among considered models. In this study, different count regression models, namely; Poisson, negative binomial, zero-inflated Poisson and zero-inflated negative binomial models were considered. Different model selection criteria: the Log likelihood, Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used in order to identify the most appropriate fitted model. In cases of over-dispersion, the ZIP model typically fits better than a standard Poisson model. But there are other models that allow for over dispersion: the standard negative binomial regression model and ZINB model. Since the Poisson and NB are not nested within the ZIP and ZINB models, respectively, we use the Vuong test to compare the non-nested models: ZIP versus Poisson and ZINB versus NB regression models.

Table4.7. Model selection criteria for Poisson, NB, ZIP and ZINB models for the number of migrants.

	Models			
Selection criteria	**Poisson**	**NB**	**ZIP**	**ZINB**
Log likelihood	-1104.486	-1082.340	-573.414	-1019.638
AIC	2258.334	2212.681	1198.828	2089.276
BIC	2369.152	2323.498	1318.845	2204.711
Voung			6.023	9.801
p-value			0.000	0.000

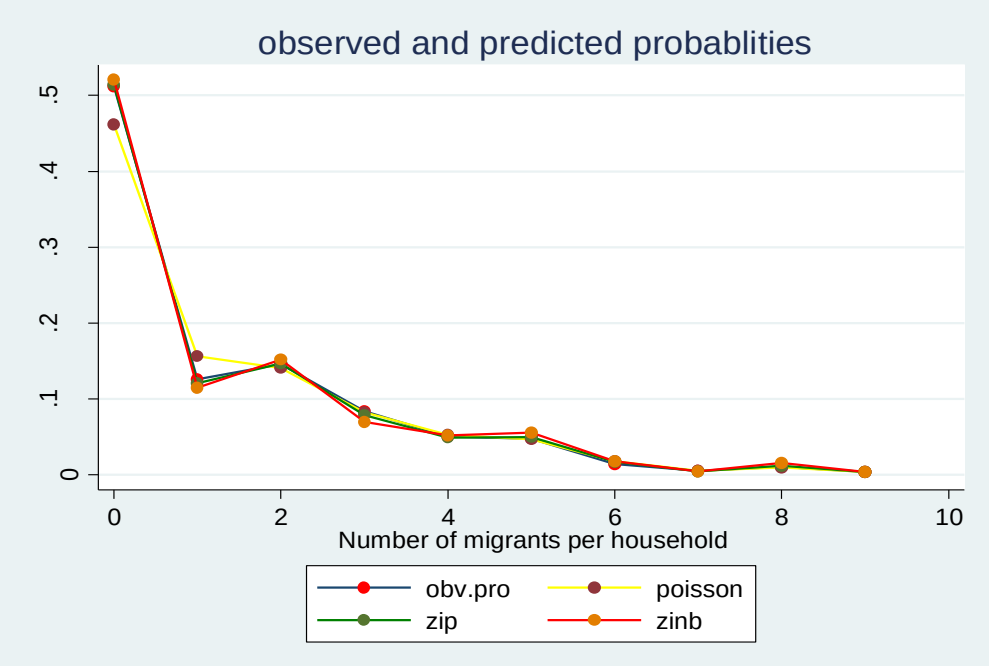
Table 4.7 shows the model selection criteria used to identify the best/preferred model among the candidate models. To compare models, AIC, BIC and Log likelihood were used as shown in Table 4.7. The model with the smallest AIC, smallest BIC and largest Log likelihood is preferred. Since ZIP model has the smallest AIC, smallest BIC and maximum Log likelihood, ZIP model is the most appropriate and preferred model among the models. Thus, the Zero-inflated Poisson regression model with the lowest value of AIC, lowest value of BIC and the highest value of Log likelihood is the most appropriate and preferred model for describing the number of migrants per household.

Table4.8 Observed and predicted probabilities of number of migrants per household

			Probabilities			
Count	observations	Observed probability	Poisson	NB	ZIP	ZINB
0	386	0.512	0.462	0.467	0.514	0.521
1	95	0.126	0.156	0.151	0.121	0.125
2	110	0.146	0.141	0.145	0.147	0.152
3	63	0.084	0.082	0.085	0.079	0.071
4	38	0.050	0.053	0.042	0.049	0.052
5	36	0.048	0.047	0.048	0.051	0.056
6	11	0.015	0.018	0.012	0.017	0.018
7	4	0.005	0.005	0.004	0.005	0.005
8	8	0.011	0.011	0.014	0.012	0.016
9	3	0.004	0.004	0.004	0.004	0.004

Table4.8indicates that predicted probabilities for ZIP model were the closest to the observed probabilities, the ZIP model is the most appropriate and preferred model than the other count models.

Figure4.4: The Observed and predicted probabilities of number of migrants per household



Since the predicted probabilities for ZIP model were the closest to the observed probabilities, the ZIP model is the most appropriate and preferred model than the other count models as shown above.

Table4.9. The estimated Zero-inflated Poisson model for number of migrants of selected independent variables.

totmigrant	Coef.	**Robust Std. Err.**	**z**	**IRR**	**P>z**	**[95% CI for Coef]**	
sex HHH							
Male	0.3236	0.0843	3.8400	1.3821	0.0000	0.1584	0.4888
Age HHH							
45-60yr	0.0720	0.0037	2.002	1.0720	0.0010	0.0001	0.0844
>60yr	0.0540	0.1600	2.169	1.0540	0.0000	0.0190	0.1025
HHH relig.							
Orthodox	-0.0685	0.1024	-1.9700	0.9338	0.0004	-0.2692	-0.0322
Muslim	-0.3696	0.1725	2.1400	0.6910	0.0020	-0.7014	-0.0377
other	-1.6909	0.3092	5.4700	0.1843	0.0000	-2.2970	-1.0849
HH size	0.1060	0.0153	6.9400	1.1118	0.0000	0.0760	0.1360
HH homeownership							
by rent	0.3988	0.0996	4.0100	1.4923	0.0000	0.2037	0.5940
other	-0.7384	0.1148	6.4300	0.4776	0.0000	-0.9634	-0.5135
Ethnicity							
Kembata	-0.0286	0.1952	-2.1500	0.9718	0.0010	-0.3539	-0.0125
Silte	-0.8251	0.1927	-4.2800	0.4382	0.0000	-1.2029	-0.4474
other	-0.1055	0.1256	-2.8400	0.8999	0.0010	-0.3517	-0.0407
Place of residence							
urban	-0.2180	0.0730	2.9900	0.8040	0.0000	-0.3050	-0.0150
job HH							
Agriculture	0.0694	0.0871	2.8000	1.0718	0.0260	0.0140	0.2402
Admn/clerical	-0.5100	0.0973	-5.2400	0.6005	0.0000	-0.7008	-0.3192
other/student	-1.5227	0.4700	-3.2400	0.2181	0.0010	-2.4439	-0.6014
Land size	-0.0503	0.0363	-2.3900	0.9509	0.0000	-0.1215	-0.0208
Mig.exp							
Migrated	0.3535	0.0842	4.2000	1.4241	0.0000	0.1886	0.5184
Education level HH							
primary	-0.0457	0.1205	-2.3800	0.9553	0.0040	-0.2819	-0.0304
secondary	-0.0489	0.1219	-2.4000	0.9523	0.0080	-0.2878	-0.0300
higher	-0.9084	0.1648	5.5100	0.4031	0.0000	-1.5855	-0.7313
ILL.ratio	0.3303	0.0663	4.9800	1.3914	0.0000	0.2004	0.4602
Dep.ratio	0.0154	0.0494	2.1100	1.0150	0.0040	0.0021	0.0914

_cons	-0.9828 1.000	0.2045 (exposure)	-4.8100	0.3743	0.0000	-1.3836	-0.5820
inflate							
HHH religion							
Orthodox	-0.5422	0.4361	-1.2400	0.5814	0.0540	-1.3970	0.3127
Muslim	4.5957	0.7765	5.9200	0.0101	0.0000	3.0737	6.1176
other	2.2272	1.3991	1.5900	0.1078	0.0623	-0.5151	4.9694
Homeownership							
by rent	-2.1926	0.7748	-2.8300	8.9800	0.0050	-3.7111	-0.6740
other	-1.2041	0.6944	1.7300	1.2041	0.0030	-0.1568	2.5651
Ethnicity							
Kembata	3.6041	1.1803	3.0500	3.6041	0.0020	1.2907	5.9175
Silte	-16.2982	1.4450	-11.2800	0.0000	0.0000	-19.1304	-13.4661
other	-2.6172	0.5836	-4.4800	0.1006	0.0000	-3.7610	-1.4733
Place of residence							
urban	-0.0318	0.3811	0.0800	0.9680	0.9330	-0.7151	0.7788
Job HHH							
Agriculture	-3.4265	0.8180	-4.1900	0.0325	0.0000	-5.0297	-1.8232
Admn/clerical	0.4536	0.3584	1.2700	0.4536	0.2060	-0.2489	1.1560
other/student	-0.7516	1.2154	-0.6200	0.4716	0.5360	-3.1338	1.6305
Mig.exp							
Migrated	-39.7142	0.8363	-47.4900	0.0000	0.0000	-41.3533	-38.0751
Education level HH							
primary	-0.6619	0.4332	-1.5300	0.5158	0.1260	-1.5109	0.1870
secondary	0.1653	0.4213	0.3900	1.1797	0.6950	-0.6604	0.9909
higher	0.4714	0.4950	0.9500	1.6022	0.3410	-0.4988	1.4416
Sex HHH							
Male	0.0258	0.3983	0.0600	1.0261	0.9480	-0.7548	0.8065
Age of HH							
45-60yr	0.6619	0.4332	-1.5300	1.9384	0.1260	-0.6950	0.0016
>60yr	0.1653	0.4950	0.3900	1.1797	0.9650	-0.5000	002360
_cons	0.0202	0.4636	0.0400	1.0202	0.9650	-0.8885	0.9289

P-value<0.05 is statistically significant. RC(reference category)=the category coded by zero

4.4.3 Interpretation of ZIP model fit results

In Table 4.8, estimated zero inflated Poisson regression model fit results of the coefficients can be interpreted as follows: for a one unit change in the predictor variable, the log of the response variable is expected to change by the value of the regression coefficient (coef.).In ZIP model, for everyone unit increase in a unit of the significant predictors, the log number of response variable is expected to increase or decrease by approximately the corresponding coefficient in the column of coefficient (coef.).In this model the variables whose p-value<0.05, were considered statistically significant. To interpret the categorical data we use the incidence rate ratio which is important to explain the change in percentage of significant predictors.

Table 4.8 contains coefficients for the factor change in the expected count for those in the Not Always-0 Group. The coefficients can be interpreted in the same way as usual.

Sex of HHH: Number of migrants in the household does differ by sex of household head. The coefficient for Male is positive and statistically significant. The positive estimated coefficient, 0.3236for *male headed household* suggests that Male headed households are expected to have about 38 percent more number of migrants than female headed household.

HHH Religion: The estimated coefficients for all religion groups of household head are negative and statistically significant. The results in Table 4.8 show that religion category of household head has a significant impact on the number of migrants in the household. That is, households with Orthodox household head are expected to have 7% percent fewer number of migrants than households with protestant households heads. Similarly, Households with Muslim household head are expected to have about 31% fewer desired number of migrants than households with protestant households heads holding all other variables in the model constant.

Home ownership: The coefficients for the categories of home ownership: own the home by rent is positive and statistically significant. The result intable4.8 implies that the household living in rented house are expected to have about49percentmore expected number of migrants than the households who living in their own house and the household who living in the home by other type of home ownership are expected to have 52 percent less expected number of migrants than the households who living in their own house.

Ethnicity: Number of migrants does differ by ethnicity of household. The coefficient for Silte ethnicity is negative and statistically significant implying that the numbers of migrants in the household that who were Silte ethnic group are expected to have about 56 percent less number of migrants than the households who are ethnic group of Hadiya. Similarly, the households who are Kembata ethnic group are expected to have about 3 percent less number of migrants than the households who are ethnic group of Hadiya

Place of Residence: Number of migrants does differ by place of residence. The coefficient for *urban* is negative and statistically significant implying that the expected number of migrant that urban households are expected to have 20% less expected number of migrants than the rural households.

Previous migration experience of household head: Number of migrants does differ by migration experience of house hold head. The coefficient for migrated household head is positive and statistically significant implying that number of migrants in households of migrated household head are expected to have about 42% more expected number of migrants than households of non migrated household head.

HHH Education: In the fitted model, the coefficient for *secondary education* is negative and statistically significant. Thus, households in which secondary level educated household head are expected to have about 5% percent fewer numbers of migrants than the household from uneducated household head. Similarly, the coefficients for both *primary* and *higher* education are negative and statistically significant. The results reveal that the household in which primary and higher educated household heads; are expected to haveabout4 percent and about 60 percent respectively fewer expected number of migrants than uneducated household head controlling for other variables in the model.

Size of house hold member: The coefficient for house hold size is positive and statistically significant implying that the expected number of migrants in the households was multiplied by1.105 for one unit increment of house hold size. In other words, household with small number of household member would likely to have smaller number of migrants in the household.

Size of land: The coefficient for house hold land size is negative and statistically significant implying that the expected number of migrants in the households was multiplied by-0.95 for one

unit increment of house hold land size. In other words, household with small size of land would likely to have large migrants in the household.

Job occupation: The estimated coefficient for household heads occupied in agriculture is positive and statistically significant**.** The positive estimated coefficient, 0.069 for *Agriculture* suggests that households rely on agriculture are expected to have about 7% percent more expected number of migrants than households rely on trade. Similarly, households rely on clerical job are expected to have about 40 percent less expected number of migrants than households rely on trade.

Dependency ratio: The coefficient for dependency ratio is positive and statistically significant implying that the expected number of migrants in the households was multiplied by0.015 for one unit increment of dependency ratio.

Illiteracy Ratio: The coefficient for illiteracy ratio is positive and statistically significant implying that the expected number of migrants in the households was multiplied by0.33 for one unit increment of illiteracy ratio.

Another part of Table 4.8, labeled "inflate", contains coefficients for the factor change in the odds of being in the Always-0 Group compared to the NotAlways-0 Group. As shown in Table 4.8, Migration experience has a significant impact on the probability of being in the always zero group. The odds of being in the always zero group decreased by about 3% for urban households as compared to those household in rural centers controlling other variables in the model. Similarly, the odds of being in the always zero group increased by about 1%for Muslim household head as compared to Orthodox household head holding all other variables in the model constant. Conversely, the odds of being in the always zero group decreased by about 7% for the household rely on agriculture as compared to household rely on trade holding all other variables in the model constant.

4.4.4 Discussion

This study tried to achieve factors that affect number of migrants in household and its consequence in Soro and Misha districts, Hadiya Zone by using count regression model. Sex of HHH, age of HHH, land size, educational, family size, place of residence and occupation were found to be important determinants of migration. Push factor, pull factor, environmental cause, positive consequence, negative consequence, economical change, and problem on the journey were also identified.

Based on the study result, sex of household head was a significant factor for number of migrants per household. The highest number of migrants per household occurred in the male headed household. This finding is consistent with other studies **(**AbrahamW. 2015).This study reveals that households with male heads can take care of the household members and manages the household chore by letting other members migrate for work. Hence, male headed households were high probability to migrate than female headed household.

Occupation of the household head has significant effect on migration. Household heads occupied in agriculture have relatively high number of migrants than household's heads rely on trade. This finding seemed to be in accordance with other studies (AbrahamW. 2015). Another research cited that the probability to be migrated for people in trade was higher than probability to be migrated for people engaged in agriculture (Teshomeet.al (2013))

According to the results, Education of the household head significantly affects number of migrants in household. Our study reveals that household with educated household head were fewer number of migrants than uneducated household head. This result was similar to the study (Taylor, J. E. 1999).

The result of this study also showed that land size significantly affects number of migrants per household. Household with small size of land have high migrants in the household. It was identified that land holding can influence number of migrants per. The size of land holding has a negative relation with number of migrants. This finding was consistent with other studies (Mckenzie, D. and H. Rapoport (2007)). From analysis of the model, it was found that family size significantly affects number of migrants per household. This finding was consistent with other studies (Tsedeke L. and Ayele T., 2017).

The study revealed that flow of peoples to RSA became noticeable beginning from1990E.C.It can easily be observed that the number of migrants is increasing overtime. This finding was similar with other studies (Teshome et.al (2013)).

In our study, family size of household head was a significant factor for number of migrants per household. A unit increases in family size of the household increases the log odds of the number of the migrant in household. This result coincides with other findings (Teshome et.al (2013)).

Place of current residence is important variables which help us to see how residence is associated with the migration. The variables are found to have a significant association with number of migrants per household. Our study revealed that the expected number of migrant that urban households are expected to have less expected number of migrants than the rural households. This finding is consistent with other relate study (Teshomeet.al (2013)).

According to the result of this study, educational level of migrants; illiterates which were about 4.9% were the lowest of all education status. From out of all 961 migrants, about 4.9% were illiterate about 16.9% were attended primary level, about 70.9% of migrants were attended secondary and about 6.9% of migrants attended higher level education. This result indicates that the most exposed by illegal migration educational level is secondary education level (Teshome D,2013).

The Major Causes of Illegal Migration of the study area was push factors of young adult migration; Poverty has contributed the highest proportion followed by Unemployment, family pressure, and peer pressure. Unemployment is series problem in this area by the Oberg, S. (2012) also asserted "For rural people, migration is one of several coping strategies to deal with poverty which in itself reflects a combination of social, economic and political conditions. Unemployment rates are usually high, while the fact that in most of these countries their earlier political and ideological regimes have collapsed intensifies and aggravates this phenomenon.

CHAPTER FIVE
5. Conclusions and Recommendations
5.1. Conclusions

This study found that some of the demographic and socioeconomic variables have a significant influence on migration to Republic of South Africa. The investigation was done mainly on the quantitative and qualitative data collected via questionnaire from 754 randomly selected households in Soro and Misha districts between March and April 2018. Information about migrants is gained from their families at home land that is proxy sampling method. In doing so, the key research questions set to be answered were: (i) Are their differences in the socioeconomic and demographic characteristics between migrant and non-migrant households? (ii) What factors initiate peoples to migrant? (iii) What are the socioeconomic and demographic consequences of such migration on the community?

This paper attempted to identify and analyze the determinants of the number of migrants per households, using count regression model. The Zero-inflated Poisson regression model (ZIP) was found to be the most appropriate and preferred model among the count regression models considered. The descriptive results suggested that there is high variability in the non-zero values. The variance of the number of migrants was larger than its mean, suggesting the possibility of over-dispersion. In addition, the over-dispersion parameter alpha was found to be significantly different from zero in Negative binomial regression models. This study captured predictor variables that had significant effects on the number of migrants. The selected ZIP model fit results indicated that sex of household head, level of education of house hold head, religion of household head, household size, home ownership, place of residence and land size were statistically significant factors influencing the number of migrants that households would like to have. The two major causes of illegal migration considered are push and pull factors. From push factor, Poverty has contributed the highest percentage followed by unemployment family pressure, peer pressure.

5.2. Recommendations

After analyzing the main causes that has high contribution to migration and consequences of labor power migration from the Soro and Misha districts of Hadiya zones to the RSA, the researcher proposes the following suggestions that could be implemented by policy makers, government official), International Organizations, and leaders or preachers of the religions as well as the community at large in the study area. Based on the findings that we have obtained, we recommend the following issues

- The government, concerned institutions and other involved stake holders hold consider the identified major factors while designing policy that will impact migration decisions the most.
- The government has to consider building vocational training institutions and agro industrial firms to create various job opportunities for community.
- To change the attitude of the community people as well as the stakeholders or preachers of the religions has to be reshaped by continuous public discussions and awareness creation programs.
- The governmental and the non-governmental organizations working in the districts should work intensively to raise the level of awareness of the whole community on migration.
- Our results justified the importance of awareness creation through training on the negative consequences of migration for the households in general and migrants in particular. It is also highly recommended that further in-depth research be conducted with respect to the causes of migration abroad, its consequences and impact on the family, challenges and coping mechanisms, multidimensionality of the migration process, and the security of migrants and livelihood of those who stay away from their homes.
- A further extension of the ideas in this study is to consider other count regression model like hurdle model, and zero-inflated generalized Poisson regression model.

REFERENCES

Abdu S. (2009). The Socio-Economic and Demographic Characteristics of Ethiopian Emigrants to the Middle East: Through Five Selected Private Employment Agencies. *Unpublished M.Sc Thesis, Addis Ababa University: Addis Ababa.*

Abraham W. (2015). Determinants of migration and its contribution on household income in Raya Azebo; Tigrai, Ethiopia.

Akaike, H. (1973). "*Information Theory and an Extension of the Maximum Likelihood Principle*." Pp. 267–281 in Second International Symposium on Information Theory, edited by B. N. Petrov and F. Csaki. Budapest: Akademiai Kaido.

Agresti, A. (2002). *Categorical Data Analysis.* 2nd edn, John Wiley and Sons, New York.

Agresti, A.(2007). *An introduction to categorical data analysis. second edition*. John Wiley and Son Inc., New York.

Bereket et.al (2011). Hazards and and associated factors among returned migrants living at Bati Woreda, Amhara National Regional State.

Berhanu, B. & White, M. (2000). War, Famine, and Female Migration in Ethiopia, 1960-1989. *Economic Development and Cultural Change, 49(1): 91-113*.

Bhargava, A., Docquier, F. & Moullan. Y. (2011). 'Modeling the Effects of Physician Emigration on Human Development.' *Economics and Human Biology 9(2): 172-83.*

Cameron A.C. and Trivedi P.K. (1998). Regression Analysis of Count Data. Cambridge University Press.

Casacchia, O, Crisci, M, & Reynaud, C. (2001). Internal migration in Ethiopia. In: Migration and urbanization in Ethiopia, with special reference to Addis Ababa. Central Statistical Authority (CSA)/Institute for Population Research – National Research Council (Irp-Cnr), October, Addis Ababa/Rome.

Clark G. and Valerie M. (2012). Drought and population mobility in rural area. *World Development,40(1), 134-145*

Clemens, M. A., & Pettersson, G. (2006). A New Database of Health Professional Emigration from Africa, Working Paper, 95 Centres for Global Development.

Cochran, W. G. (1977). *Sampling Techniques. 3rd Edition*. Harvard University, New York.

Collett, D.(2003). *Modeling Binary Data, 2nd edition, Chapman and Hall*, London, UK.

Dejene Aredo, (2011). Migrant Remittances Shocks and Poverty in Urban Ethiopia: *An Analysis of Micro Level Panel Data. Addis Ababa University*

Emerta Assaminew et al., (2010). "International Migration, Remittances and Poverty Alleviation in Ethiopia:" *Ethiopian Economics Association (EEA) / Ethiopian Economic Policy Research Institute (EEPRI), working paper series, No. 1 /2010, Addis Ababa, Ethiopia*

Fajnzylber, P. and J. Humberto Lopez (2007). The Development Impact of Remittance in Latin America. Washington DC: the World Bank

Girum S. (2010). 'Hiywot yemiyaskefilewu ye debub Africa ye sidet guzo' *Admass Advertising [Life Costing Migration to South Africa. Addis Admass Magazine Vol 10 Number 552, August 13, 2010]*

Harris, J., & Todaro, M. (1970). Migration, unemployment and development: a two-sector analysis. American Economic Review, 60, 126–142

Hilbe, J. M. (2011). *Negative binomial regression*, second edition, Cambridge University press.

International Centre for Migration Policy Development (ICMPD) (2008). East Africa Migration Route Initiative Gaps & Needs Analysis Project Country Reports: Ethiopia, Kenya, and Libya.

IOM (2004). Glossary of International Migration Law. Geneva: International Organization for Migration. Retrieved on 24 October 2017

Lee, Everett S. 1968. "Theory of Migration", D. M.Heer (Eds.) *Readings on Population*. New Jeresy: Prenctive Hall

Lee E. (1997). "The theory of Migration." *Journal of the Royal statistical Society.*

Lindsey, J.K. (1999). On the use of corrections for over-dispersion, Applied Statistics, 48(4):553-561.

Messay M. (2005). 'Debub Africa: ye tesfayitu midir' Addis Admass Gazeta Qitse 5 Hidar 13 to Yekatit 30, 1997 E.C Hibreteseb Amid page-8 Addis Ababa: Admass Advertising [South Africa: the Dream Land. Addis Admass Magazine November 22 to March 10, 2005 page 8, Addis Ababa: Admass Advertising]

Massey et al. (2011). Theories of international migration: a review and appraisal," Population and Development Review 19 (3): 431-466.

Massey, D.S. and K.E. Espinosa (1997). "What's Driving Mexico-U.S. Migration? A Theoretical, Empirical, and Policy Analysis",*The American Journal of Sociology*, 102, No.4, pp. 939-999.

Mccullagh,P and Nelder J.A, (1989). Generalized Linear Models, Monographs on Statistics and Applied Probability 37, 2nd edn, Chapman and Hall, London,UK.

Mckenzie, D. and H. Rapoport (2007). "Network effects and the dynamics of migration and Inequality: Theory and evidence from Mexico", Journal of Development Economics, 84, Issue 1, pp. 1-24

Migration and Globalization (2010). A project of the Levin Institute (The State University of New York),

Nigatu Regassa and Ansha Yusufe (2007). Journal of Social and Economic Development, 2007, vol. 9, issue 1, pages 25-38.

Nwokocha, E.E (2007). Gender inequality and Development in Nigeria: A Review on Antithesis South-South Journal of Culture and Development. Forth coming.

Oberg, S. (2012), Spatial economic factors in future South-North Migration: W. Lutz (ed.), The Future population of the world: What can us Assume Today? Earth scan, London: 336- 357

Sanne Schroll, (2008). "Return Migration - A Duration Analysis" PRELIMINARY VERSION September 1, 2008.

Sindu (2007). Irregular Migration Flows in the Horn of Africa Challenges and implication for source, transit and destinations countries.

Taylor, J. E. (1999). The New Economics of Labour Migration and the Role of Remittances in the Migration Process. International Migration, 37 (1), 63-88.

Taylor, J. (1986). Differential migration networks, information and risk. In O. Stark (Ed.), Research in human capital and development (Vol. 4). Greenwich, CN: JAI Press

Tsedeke L. and Ayele. (2017). Logistic Mixed Modeling of Determinants of International Migration from the Southern Ethiopia: Small Area Estimation Approach. *American Journal of Theoretical and Applied Statistics*
2017; 6(3): 170-182.

Terrazas, A.M, (2007). Beyond Regional Circularity: The Emergence of an Ethiopian Diaspora. Available at: http://www.irinnews.org/report.aspx?reportid=49412 (accessed on 20 December 2018).

Teshome et.al (2013). Irregular Migration: Causes and Consequences of Young Adult Migration from Southern Ethiopia to South Africa

UNDESA (2017). Trends in International Migrant Stock: Migrants by Destination and Origin (United Nations database, POP/DB/MIG/Stock/Rev.2013).

United Nations Development Programme (UNDP) (2010). *Human Development Report (2009). Overcoming Barriers: Human Mobility and Development.* New York: UNDP.

United Nation statistics Division (UNSD) (2017). Gender, Demographic and Social statistics. Available at :http://unstats.un.org/home/ (accessed on 14 March 2018).

World Migration (2014). Managing Labour Mobility in the Evolving Global Economy, Volume 2, International Organization for Migration, Hammersmith Press: IOM.

World Bank (2014). Migration and Remittances Factbook 2011, Washington DC: The World Bank.

Addis Ababa University, College of Natural and Computational Science

Department of Statistics *(Graduate program (R))*

Dear respondents: *Thank you very much for your time!* The purpose of this questionnaire is to gather information to conduct research for the partial fulfillment of the requirements for the degree of Master of Applied Statistics. This instrument is designed for the purpose of gathering information concerning the determinants of migration and its related impacts of migration in Soro and Misha woreda (district) to the Republic of South Africa (RSA). The information that you will have provide is intended to serve for identification of determinant factors of migration and its positive and negative consequences in origin country as well as its implications for bilateral relations of Ethiopia and RSA. Therefore, you are kindly requested to provide accurate information as much as possible. I confirm you that all data given by you will be treated confidentially.

N .B For this research, out migrants include the migrants who prepared themselves to go RSA and who have gone to RSA.

Instruction: Thick in the box or write the answer as may be necessary and Do not write your name

I would like to Thank you again very much for your time!

1. Is there any migrant (Return migrant or Out migrant) to RSA in your household?
   - ☐ Yes
   - ☐ No
2. If Q1 answer is yes, how many *total* migrants (return migrants and out migrants) are present in your household? --------
3. If Q1 answer is yes, how many *female* migrants (return migrants and out migrants) are present in your household? --------
4. If Q1 answer is yes, how many *male* migrants (return migrants and out migrants) are present in your household? --------
5. If Q1 answer is yes, List the number of migrants in your household by their demographic characteristics like age, sex, educational level, etc .. in the following table (thick in the box)

Migrants	Sex	Age	Educational level	year of migration to RSA(year in E.C)	previous migration status of migrants(except to RSA)	economic activity

Migrant1	□ Male □ Female	□ <15 □ 15-30 □ 31- 45 □ >45	□ Illiterate □ Primary □ Secondary □ Higher		□ migrant □ non-migrant	□ Agriculture □ Merchant □ Admins/clerical □ Other
Migrant2	□ Male □ Female	□ <15 □ 15-30 □ 31- 45 □ >45	□ Illiterate □ Primary □ Secondary □ Higher		□ Migrant □ non-migrant	□ Agriculture □ Merchant □ Admins/clerical □ Other
Migrant3	□ Male □ Female	□ <15 □ 15-30 □ 31- 45 □ >45	□ Illiterate □ Primary □ Secondary □ Higher		□ Migrant □ non-migrant	□ Agriculture □ Merchant □ Admins/clerical □ Other
Migrant4	□ Male □ Female	□ <15 □ 15-30 □ 31- 45 □ >45	□ Illiterate □ Primary □ Secondary □ Higher		□ Migrant □ non-migrant	□ Agriculture □ Merchant □ Admins/clerical □ Other
Migrant5	□ Male □ Female	□ <15 □ 15-30 □ 31- 45 □ >45	□ Illiterate □ Primary □ Secondary □ Higher		□ Migrant □ non-migrant	□ Agriculture □ Merchant □ Admins/clerical □ Other
list if exist accordingly • . •	• (…)	• (…)	• (…)	(…)	(…)	(…)

6. Sex of household head

   □ Male Female

7. Age of household head_____

8. What is your household heads religion?

- Protestant
- Orthodox
- Muslim
- other

9. House hold member size_____
10. Households home ownership
    - Own
    - By rent
    - Other
11. Household heads ethnicity
    - Hadiya
    - Kembata
    - Silte
    - Other
12. Current residence place of the household.
    - Rural
    - urban
13. Main economic activity or job of the house hold head
    - Agriculture
    - Merchant
    - Admin/clerical
    - Other
14. Land size for household in hectare____
15. Does your household head experienced migration to else abroad country except RSA?
    - Yes
    - No
16. What is highest level of education for HHH?
    - Illiterate
    - Primary (1-8)
    - Secondary (9-12)
    - Higher
17. How many illiterates are in your house hold? _____
18. How many members of household are aged less than 15 years? _____
19. How many members of household are aged greater than 65 years? _____
20. How much is your households' average monthly income now in ETB*?(Not that: it should not include the income remitted from RSA!)* _____

Questionnaire for the Causes of Migration

21. What are the "push" factors which motivate people to migrate to the RSA in your Environment? (It is possible to choose more than one response)
    - Poverty
    - Unemployment
    - Family pressure
    - Peer pressure
    - Population growth
    - Land shortage
    - Large family size
22. What are the "pull" factors do you think that motivate people to migrate to the RSA in your Environment? (You can choose more than one response).
    - High income in RSA
    - Social networks
    - Job opportunity
    - Smugglers

Questionnaire for the Consequences of Migration

23. If your answer for Q1 is yes is there any difference before and after migration in your economy or livelihood? Or is migration has changed your economy? A. Yes B. No C. Not sure
24. Does migration to RSA has positive consequence at your Environment? A. Yes B. No C. Not Sure
25. If you say "yes" for Q.24, what type of positive consequence has migration to RSA in your Environment? (You can circle more than one response)
    - Flow of remittance
    - Job creation opportunities
    - Diaspora benefits
    - Poverty has decreased
    - Improvement of social services
    - Other (specify
26. Does migration to RSA has negative consequence at your Environment? A. Yes B. No C. Not Sure
27. If you say "yes" for Q.26 which one of the following could be the negative consequences of migration to RSA in your woreda/ Environment?
    - Economic Inequality among the people

- Dependency on remittance
- Shortage of labour force
- Lack of job creation interest
- Loss of life
- Others specify _____

28. Does migration to RSA has negative consequences on students learning or education? A. Yes B. No C. Not sure
29. If you say "yes" for Q.26, which one of the following are negative impacts of migration to RSA on learning? (Please circle more than one response)
    - School dropouts
    - Brain drain
    - Absenteeism
    - Low achievement
30. In which way the majority people have been migrating to the destination country or RSA? A. In regular way B. In irregular way
31. If your answer for Q30 is in irregular way, what do you think that why do they choose this way?
    - Because it is Less cost
    - To consume lesss time
    - Has no any protection
    - To reduce bureaucracy
32. What problems do you think that have been facing when journey to RSA in irregular way? (It is possible to choose more than one response)
    - Imprisonment
    - Robbery
    - Human trafficking
    - Beaten by police
    - Death
    - Lack of food and water
    - Lack of residence

Thank you!!!!!

Appendices-1

TableA1.1: Percentage of respondents by Major Causes and consequence of Migration

		Freq.	percentage
Change in economy	*no*	36	10%
	yes	318	86%
	not sure	14	4%
Positives Consequence	*no*	90	12%
	yes	648	86%
	not sure	18	2%
Type of positive consequence	*remittance*	617	95%
	job creation	42	6%
	Diaspora benefit	94	15%
	poverty decreased	638	98%
	improvement in social service	181	28%
	other	17	3%
negative consequence	*no*	178	24%
	yes	558	74%
	not sure	18	2%
type of negative consequence	*economic inequality*	358	64%
	depending on remittance	189	34%
	shortage of labor force	265	47%
	lack of job creation	158	28%
	loss of life	549	98%
	other	73	13%
students consequence	*no*	148	20%
	yes	452	60%
	not sure	154	20%
Type of students consequence	*dropouts*	312	69%

	Brain drain	22	5%
	Absenteeism	367	81%
	low achievement	128	28%
Migration way	*irregular way*	629	83%
	regular way	125	17%
Reason irregular way			
	less cost	603	96%
	less time	135	21%
	no protection	33	5%
	reduce bureaucracy	16	3%
Problem irregular way			
	Imprisonment	108	14%
	Robbery	259	34%
	Human trafficking	57	8%
	Beaten by police	82	11%
	Death	467	62%
	Lack of food and water	536	71%
	Lack of residence	369	49%

Source: field survey (March- April 2018)

Appendices-2

TableA2.1. Estimates and standard errors for Poisson model

```
Poisson regression                              Number of obs   =        754
                                                Wald chi2(24)   =     447.83
                                                Prob> chi2      =     0.0000
```

Log pseudolikelihood =-1104.486

totmigrant	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
1.sexhhh						
male	0.323	0.113	4.23	0.000	-0.2355814	0.641894
agehhh						
45-60year	0.077	0.174	1.8	0.012	-0.0006071	0.0141896
>60year	0.054	0.138	2.3	0.002	-0.0005600	0.0145870
hhhrelg						
1	-0.062	0.114	-0.05	0.003	-0.2200171	0.2095081
2	-0.387	0.289	-2.56	0.001	-1.064249	-0.141925
3	1.690	0.336	3.24	0.000	-0.3937121	1.597862
hhsize	0.100	0.018	-0.44	0.000	-0.0467427	0.0296444
hhomeownership						
1	0.196	0.136	0.64	0.002	-0.1416695	0.2802877
2	0.734	0.131	4.12	0.000	-0.2873323	0.8089093
ethincity						
1	-0.029	0.196	-3.57	0.000	-0.939988	-0.274006
2	-0.681	0.194	-2.83	0.005	-0.8821159	-0.159825
3	0.089	0.137	1.22	0.002	-0.0877371	0.3781177
1.rphh	-0.217	0.092	2.63	0.009	-0.059596	0.4080929
jobhh						
1	-0.067	0.119	4.1	0.000	0.226207	0.640044
2	-0.495	0.126	-3.32	0.001	-0.6124005	-0.157938
3	-1.661	0.503	-3.24	0.001	-2.301525	-0.567583
Lsizehh	-0.08	0.032	1.86	0.063	-0.003167	0.1186581
1.hhmexp	-0.363	0.098	11.83	0.000	0.8778088	1.22642
edulevelhh						
1	-0.046	0.122	-0.07	0.004	-0.2446388	0.2289231
2	-0.047	0.119	-0.63	0.008	-0.3036748	0.156123
3	-0.034	0.172	3.43	0.001	0.2505941	0.9165801
ILratio	0.264	0.071	3.84	0.000	0.1304713	0.4024918
Dep_ratio	0.007	0.084	0.08	0.004	-0.1553193	0.1693026
incomehh	0.000	0.000	0.0065	0.099	-0.0000020	0.0000239
_cons	-1.935	0.058	-11.51	0.000	-3.599983	-2.552197
ln(hhsize)	1	(exposure)				

estatic

Akaike's information criterion and Bayesian information criterion

model	Obs	ll(null)	ll(model)	df	AIC	BIC
	754	-1273.863	-1065.486	25	2258.334	2369.152

Note: N=754 observation used in calculating BIC

estatgof

Deviance goodness-of-fit	=	1219.434
Prob> chi2(729)	=	0.000
Pearson goodness-of-fit	=	1302.256
Prob> chi2(729)	=	0.000

TableA2.2 Estimates and standard errors for NB model

Negative binomial regression Dispersion = mean

Wald chi2 (24) = 385.04 Number of obs = 754

Log pseudolikelihood = -1082.340 Prob> chi2 = 0.0000

totmigrant	ceff	std err.	z	P>z	[95% Conf.	Interval]
1.sexhhh	0.327	0.115	4.13	0.000	0.243823	0.683309
agehhh						
45-60year	0.081	0.269	1.79	0.001	-0.00075	0.016271
>60year	0.050	1.135	1.92	0.001	-0.00160	0.15920
hhhrelg						
1	-0.03	0.116	0.02	0.987	-0.23767	0.241577
2	-0.60	0.27	-2.94	0.003	-1.14891	-0.22985
3	1.152	0.366	3.19	0.001	0.396924	1.658245
hhsize	0.102	0.02	-0.19	0.852	-0.04329	0.035779
hhomeownership						
1	0.195	0.139	0.41	0.684	-0.20381	0.310572
2	0.735	0.158	2.71	0.007	0.127136	0.794285
ethincity						
1	-0.03	0.205	-3.84	0	-0.97024	-0.31446
2	-0.76	0.223	-2.60	0.009	-1.01103	-0.14165
3	0.103	0.143	0.36	0.721	-0.22653	0.327367
jobhh						
1	-0.06	0.113	4.09	0	0.264557	0.752465
2	-0.49	0.141	-2.59	0.01	-0.69339	-0.09609
3	-1.61	0.493	-3.67	0	-1.85054	-0.56304
Lsizehh	-0.25	0.103	1.58	0.114	-0.01455	0.136251
1.hhmexp	-0.32	0.105	9.06	0.000	0.859056	1.333656
edulevelhh						

1	-0.04	0.148	-0.17	0.862	-0.27612	0.231243
2	-0.04	0.138	-0.79	0.428	-0.36432	0.154453
3	-0.68	0.212	3.1	0.002	0.208112	0.926558
ILratio	0.358	0.093	2.77	0.006	0.074922	0.438360
Dep_ratio	0.016	0.082	-0.13	0.897	-0.18551	0.162519
incomehh	0.00	0.00	1.71	0.087	-2.2E-06	3.25E-05
_cons	-2.45	0.003	-10.5	0.000	-3.77880	-2.59452
ln(hhsize)	1	(exposure)				
/lnalpha	-0.85	0.22			-1.36170	-0.46062
alpha	0.607	0.117			0.25622	0.63089

Likelihood-ratio test of alpha=0: chibar2 (01) = 36.24 Prob>=chibar2 = 0.000

---

estatic

Akaike's information criterion and Bayesian information criterion

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
	754	-1140.36	-1047.36	26	2212.681	2323.498

TableA2.3 Estimates and standard errors for ZIP model

Zero-inflated Poisson regression Number of obs =754

Nonzero obs =368

Zero obs = 386

Inflation model = logit Prob> chi2 = 0.0000

Log likelihood = -573.414

totmigrant	Coef.	**Robust Std. Err.**	**z**	**IRR**	**P >z**	**[95% Conf.**	**Interval]**
sex HHH							
Male	0.3236	0.0843	3.8400	1.3821	0.0000	0.1584	0.4888
Age HHH			2.002				
45-60yr	0.0720	0.0037	2.169	1.0720	0.0010	0.0000	0.0144
>60yr	0.0540	0.1600		1.0540	0.0000	0.0690	0.1025

HHH relig.							
Orthodox	-0.0685	0.1024	-1.9700	0.9338	0.0004	-0.2692	-0.0322
Muslim	-0.3696	0.1725	2.1400	0.6910	0.0020	-0.7014	-0.0377
other	-1.6909	0.3092	5.4700	0.1843	0.0000	-2.2970	-1.0849
HH size	0.1060	0.0153	6.9400	1.1118	0.0000	0.0760	0.1360
HH homeownership							
by rent	0.3988	0.0996	4.0100	1.4923	0.0000	0.2037	0.5940
other	-0.7384	0.1148	6.4300	0.4776	0.0000	-0.9634	-0.5135
Ethnicity							
Kembata	-0.0286	0.1952	-2.1500	0.9718	0.0010	-0.3539	-0.0125
Silte	-0.8251	0.1927	-4.2800	0.4382	0.0000	-1.2029	-0.4474
other	-0.1055	0.1256	-2.8400	0.8999	0.0010	-0.3517	-0.1407
Place of residence							
urban	-0.2180	0.0730	2.9900	0.804	0.0000	0.0750	0.3610
job HH							
Agriculture	0.0694	0.0871	2.8000	1.0718	0.0260	0.0140	0.2402
Admn/clerical	-0.5100	0.0973	-5.2400	0.6005	0.0000	-0.7008	-0.3192
other/student	-1.5227	0.4700	-3.2400	0.2181	0.001	-2.4439	-0.6014
Land size	-0.0503	0.0363	-2.3900	0.9509	0.00	-0.1215	-0.0208
Mig.exp							
Migrated	0.3535	0.0842	4.2000	1.4241	0.000	0.1886	0.5184
Education level HH							
primary	-0.0457	0.1205	-2.3800	0.9553	0.004	-0.2819	-0.0304
secondary	-0.0489	0.1219	-2.4000	0.9523	0.008	-0.2878	-0.1900
higher	-0.9084	0.1648	5.5100	0.4031	0.000	-1.5855	-1.2313
ILL.ratio	0.3303	0.0663	4.9800	1.3914	0.000	0.2004	0.4602
Dep.ratio	0.0154	0.0494	2.1100	1.015	0.004	0.0021	0.0914
Income HH	0.0000	0.0000	1.1700	1.0000	0.243	0.0561	0.0000
_cons	-0.9828	0.2045 (exposur)	-4.8100	0.3743	0.000	-1.3836	-0.5820
	1.000	)					
inflate							
HHH religion							
Orthodox	-0.5422	0.4361	-1.2400	0.5814	0.054	-1.3970	0.3127
Muslim	-4.5957	0.7765	5.9200	0.0101	0.000	3.0737	6.1176
other	-2.2272	1.3991	1.5900	0.1078	0.062	-0.5151	4.9694
Home ownership							
by rent	2.1926	0.7748	-2.8300	8.9800	0.005	-3.7111	-0.6740
other	1.2041	0.6944	1.7300	1.2041	0.003	-0.1568	2.5651
Ethnicity							
Kembata	3.6041	1.1803	3.0500	3.6041	0.002	1.2907	5.9175

Silte	-16.298	1.4450	-11.28	0.0000	0.000	-19.130	-13.4661
other	-2.617	0.5836	-4.480	0.1006	0.000	-3.7610	-1.4733
Place of residence							
urban	-0.031	0.3811	0.0800	0.9680	0.9330	-0.7151	0.7788
Job HHH							
Merchant	-3.426	0.8180	-4.190	0.0325	0.000	-5.0297	-1.8232
Admn/clerical	0.4536	0.3584	1.270	0.4536	0.206	-0.2489	1.1560
other/student	-0.751	1.2154	-0.620	0.4716	0.536	-3.1338	1.6305
Mig.exp							
Migrated	-39.71	0.8363	-47.49	0.0000	0.000	-41.353	-38.0751
Education level HH							
primary	-0.661	0.4332	-1.530	0.5158	0.126	-1.5109	0.1870
secondary	0.1653	0.4213	0.3900	1.1797	0.695	-0.6604	0.9909
higher	0.4714	0.4950	0.9500	1.6022	0.3410	-0.4988	1.4416
Sex HHH							
Male	0.0258	0.3983	0.0600	1.0261	0.9480	-0.7548	0.8065
Age of HH							
45-60yr	0.6619	0.4332	-1.530	1.9384	0.1260	-0.6950	0.0016
>60yr	0.1653	0.4950	0.3900	1.1797	0.9650	-0.5000	002360
_cons	0.0202	0.4636	0.0400	1.0202	0.9650	-0.8885	0.9289

estatic

Akaike's information criterion and Bayesian information criterion

Model	Obsll(null)	ll(model)	df	AIC	BIC
.	754	-637.3667	-573.9839	49	1199.968	1318.843

Note: N=754 Obs used in calculating BIC

TableA2.4 Estimates and standard errors for ZINB model

Zero-inflated regression Number of obs = 754

negative binomial

Nonzero obs = 368

Zero obs = 386

Inflation model = logit

Log likelihood = -1019.638

Prob> chi2 = 0.000

totmigrant	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
totmigrant						
1.sexhhh	0.3436	0.0864	3.360	0.0010	0.1394	0.528813
agehhh						
<45yr	0.0710	0.1258	1.980	0.0023	-0.1250	0.2365
45-60yr	0.0504	0.2650	2.010	0.0054	-0.2564	0.2568
hhhrelg						
1	-0.0630	0.0902	0.840	0.4000	-0.1229	0.308056
2	-0.3496	0.2944	1.310	0.1890	-0.8247	0.162652
3	1.6795	0.3789	4.580	0.0000	0.9351	2.334658
hhsize	0.1102	0.0178	2.680	0.0070	0.0143	0.091906
hhomeownership						
1	0.3932	0.1016	-1.52	0.1290	-0.4257	0.054235
2	0.6915	0.1384	4.240	0.0000	0.3711	1.008956
ethincity						
1	-0.0292	0.1863	-2.68	0.0070	-0.7726	-0.11997
2	-0.7903	0.2085	-2.84	0.0040	-1.0411	-0.19171
3	0.1022	0.1161	2.140	0.0320	0.0240	0.543351
1.rphh	-0.2511	0.0774	2.830	0.0050	0.0784	0.431388
jobhh						
1	-0.6248	0.0856	3.160	0.0020	0.1266	0.540306
2	-0.5647	0.1032	-4.39	0.0000	-0.9116	-0.34909
3	-1.662	0.4989	-3.53	0.0000	-2.0216	-0.57903
Lsizehh	-0.5001	0.0358	-2.28	0.0220	-0.1738	-0.01331
1.hhmexp	-0.3591	0.0874	5.450	0.0000	0.3406	0.72274
edulevelhh						
1	-0.1191	0.1202	1.000	0.3190	-0.1097	0.337012
2	-0.0327	0.1214	0.190	0.8530	-0.2123	0.256614
3	-1.0514	0.1811	5.270	0.0000	0.6371	1.39164
ILratio	0.3424	0.0687	3.190	0.0010	0.1215	0.509094
Dep_ratio	0.0222	0.0549	0.330	0.7430	-0.1194	0.167395
incomehh	0.0000	0.0000	1.670	0.0940	0.0000	2.57E-05
_cons	-0.6186	0.0265	-4.45	0.0000	-1.7466	-0.67811

ln(lnexposure)	1	(exposure)				
inflate						
sexhhh	-0.5593	0.25749	-2.17	0.0300	-1.06404	-0.05466
agehhh	-0.0364	0.01037	-3.51	0.0056	-0.05674	-0.01607
hhhrelg	0.38742	0.14412	2.690	0.0070	0.104951	0.669905
hhmexp	-15.501	499.555	-0.03	0.9750	-994.617	963.6144
jobhh	0.05559	0.12901	0.430	0.6670	-0.19728	0.308456
hhsize	-0.2037	0.04800	-4.24	0.0000	-0.29780	-0.10963
ethincity	0.28809	0.10889	2.650	0.0080	0.074659	0.501532
hhomeownership	0.12100	0.15616	0.770	0.4380	-0.18507	0.427079
edulevelhh	0.03815	0.12599	0.300	0.7620	-0.20877	0.285121
rphh	-0.2367	0.27795	-0.85	0.3940	-0.78156	0.308024
_cons	-0.6376	0.12146	-5.25	0.0000	-0.8757	0.32568
/lnalpha	-7.6000	0.6250	-0.04	0.9720	-975.5314	940.7639
alpha	0.00051	0.0002			0.0000001	0.0000029

estatic

Akaike's information criterion and Bayesian information criterion

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
	754	-1133.15	-1011.07	37	2089.207	2204.711