2017

Three Essays In Population Process

Romeo Jacky Gansey
University of Pennsylvania, rgansey@sas.upenn.edu

Three Essays In Population Process

Abstract

This dissertation addresses multiple topics of current population processes, including an impact evaluation of malaria control on child mortality in Tanzania, a study of living arrangements among the foreign-born in the United States, and an investigation of selectivity and the choice of migration destination among African emigrants. The dissertation follows a three-chapter format. The chapters are related to one another by their common focus on policy relevant topics for the health and wellbeing of populations. The first chapter investigates the specific contribution of malaria control to improvements in the health of children under five in mainland Tanzania by exploiting the timing of scale-up malaria interventions along with the variation in malaria endemicity across the country due to ecology. The analyses are based on birth history and socioeconomic information from the 2004-2005 and 2009-2010 waves of the Tanzanian Demographic and Health Survey and epidemiological information on malaria prevalence from the National Malaria Control Program. The results suggest that, on average, malaria control interventions have helped avert approximately 17.9 deaths for every 1,000 live births between 2004 and 2010. They also point to significant improvements in children's nutritional health attributable to malaria control.

The second chapter examines living arrangements among the foreign-born in the United States by including all major sending regions of immigrants; by distinguishing between horizontal and vertical extension (coresidence within and across generations); and by accounting for the uneven geographic distribution of immigrants across the country. Drawing on data from the five percent sample of the 2001-2013 waves of the American Community Survey, the chapter shows not only large differentials in the prevalence of extension across immigrant groups, but also substantial variation in the type and predictors of extension, and the extent to which these differences with native whites are explained by socio-demographic composition and housing conditions. Overall, traditional theories of extension do a better job of explaining horizontal than vertical extension, and among relatively disadvantaged immigrant groups (i.e., Mexicans, Latin Americans (excluding Mexicans) and West Indians) than more positively selected groups (i.e., South East Asians, Canadians/Europeans, and Oceanians). African immigrants often fall in between these two extremes. The chapter also shows that accounting for immigrant concentration in more expensive housing markets explains an important share of the immigrant-native gap in extension, suggesting that previous analyses exaggerated the role of culture in explaining variation in living arrangements.

The third chapter is concerned with emigrant selectivity and the factors that shape the choice of migration destinations. Using data from the 2009 Ghana survey of the Migration from Africa to Europe (MAFE) project and data from nationally representative surveys in the United Kingdom and the United States, this chapter investigates these issues among emigrants from Ghana. The results point to positive selection among current emigrants compared to both non-migrants and return migrants. They further show that return migrants tend to have more favorable socioeconomic characteristics than non-migrants. The results also indicate selectivity among emigrants by destination such that those who migrate to a destination further away are more positively selected. However, these results fail to show any effect of wage differentials at destination in explaining the choice of migration destination among emigrants to the United Kingdom and the United States. This chapter offers three contributions to the literature on migrant selectivity. First, it focuses on selectivity among return migrants versus current emigrants and compares emigrants across destinations. Second, it assesses whether socioeconomic characteristics used at either the individual or household level influence migrant selectivity. Finally, it tests empirically two theoretical pathways (neoclassical and network theories) relating migrant characteristics to the choice of migration destinations.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Demography

First Advisor
Irma T. Elo

Keywords
Child Health, Living Arrangements, Malaria, Migration, Program Evaluation

Subject Categories
Demography, Population, and Ecology

THREE ESSAYS IN POPULATION PROCESS

Romeo Gansey

A DISSERTATION

in

Demography

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2017

Supervisor of Dissertation

Irma T. Elo, Professor of Sociology

Graduate Group Chairperson

Hans-Peter Kohler, Frederick J. Warren Professor of Demography

Dissertation Committee

Chenoa Flippen, Associate Professor of Sociology

Petra Todd, Alfred L. Cass Term Professor of Economics

Dedicated to Camille, Fieyichitan, and Kajola

ACKNOWLEDGEMENT

Dissertations are written mostly in the deep spaces of solitude. Yet, behind my solitude there is the support of the Penn community, whose frontline pillars are my dissertation committee members. I am indebted to Irma Elo, the Chair of my dissertation committee, for her invaluable mentorship from the beginning to the end of my PhD studies and the completion of my dissertation. Irma has been for me a source of inspiring scholarship and dedication during the course these past years. She has also supported me at every step of my PhD studies, whose completion would not have been possible without her help. I wish further to thank Chenoa Flippen and Petra Todd, the other members of my dissertation committee, who have generously assisted me with advice and time. They have shown patience in revising early drafts of this dissertation.

Elizabeth Frankenberg is an excellent teacher and mentor. She showed me the way of graduate studies and taught me more about research than I could possibly express in any way. Duncan Thomas has been incredibly inspiring. His early advice encouraged me to make the decision to pursue a PhD in Demography. During the course of the past two years, Samuel Preston has generously advised me.

Students are grateful to faculty and staff for very good reasons. The Population Studies Center has provided me with a healthy environment for the pursuit of my academic goals. Michel Guillot has been a warm source of encouragement during my years at Penn. Tanya Yang and Dawn Ryan have been tremendously helpful in several ways. Nykia Perez Kibler has been instrumental in teaching me the specifics of several support tasks for research.

I have the privilege to bear further debts of gratitude. It would be impossible to adequately thank my fellow PhD students and postdocs at the PSC, with whom I have had the immense luck to interact. I thank especially Ilka Vari-Lavoisier for her honesty in revising my drafts, her incredible energy, and her generosity with time. I also thank warmly Jeylan Erman and Sarah Adeyinka-Skold for their feedback and kindness. I am grateful to Luca-Maria Pesando for his deep sense of friendship.

Finally, I am thankful to my family for their love and unwavering support. From various parts of the world, my brothers Renaud and Rudy have helped me sail through stormy waves and also share joyful moments. My mother, Leocadie, has been present for me at every decisive moment of my PhD studies and life. They deserve my deep gratitude.

ABSTRACT

THREE ESSAYS IN POPULATION PROCESS

Romeo Gansey

Irma T. Elo

This dissertation addresses multiple topics of current population processes, including an impact evaluation of malaria control on child mortality in Tanzania, a study of living arrangements among the foreign-born in the United States, and an investigation of selectivity and the choice of migration destination among African emigrants. The dissertation follows a three-chapter format. The chapters are related to one another by their common focus on policy relevant topics for the health and wellbeing of populations. The first chapter investigates the specific contribution of malaria control to improvements in the health of children under five in mainland Tanzania by exploiting the timing of scale-up malaria interventions along with the variation in malaria endemicity across the country due to ecology. The analyses are based on birth history and socioeconomic information from the 2004-2005 and 2009-2010 waves of the Tanzanian Demographic and Health Survey and epidemiological information on malaria prevalence from the National Malaria Control Program. The results suggest that, on average, malaria control interventions have helped avert approximately 17.9 deaths for every 1,000 live births between 2004 and 2010. They also point to significant improvements in children's nutritional health attributable to malaria control.

The second chapter examines living arrangements among the foreign-born in the United States by including all major sending regions of immigrants; by distinguishing between horizontal and vertical extension (coresidence within and across generations);

and by accounting for the uneven geographic distribution of immigrants across the country. Drawing on data from the five percent sample of the 2001-2013 waves of the American Community Survey, the chapter shows not only large differentials in the prevalence of extension across immigrant groups, but also substantial variation in the type and predictors of extension, and the extent to which these differences with native whites are explained by socio-demographic composition and housing conditions. Overall, traditional theories of extension do a better job of explaining horizontal than vertical extension, and among relatively disadvantaged immigrant groups (i.e., Mexicans, Latin Americans (excluding Mexicans) and West Indians) than more positively selected groups (i.e., South East Asians, Canadians/Europeans, and Oceanians). African immigrants often fall in between these two extremes. The chapter also shows that accounting for immigrant concentration in more expensive housing markets explains an important share of the immigrant-native gap in extension, suggesting that previous analyses exaggerated the role of culture in explaining variation in living arrangements.

The third chapter is concerned with emigrant selectivity and the factors that shape the choice of migration destinations. Using data from the 2009 Ghana survey of the Migration from Africa to Europe (MAFE) project and data from nationally representative surveys in the United Kingdom and the United States, this chapter investigates these issues among emigrants from Ghana. The results point to positive selection among current emigrants compared to both non-migrants and return migrants. They further show that return migrants tend to have more favorable socioeconomic characteristics than non-migrants. The results also indicate selectivity among emigrants by destination such that those who migrate to a destination further away are more positively selected. However, these results fail to show any effect of wage differentials at destination in explaining the choice of migration destination among emigrants to the United Kingdom and the United States. This chapter offers three contributions to the literature on migrant selectivity. First, it focuses on selectivity among return migrants versus current emigrants and compares emigrants across destinations. Second, it assesses whether socioeconomic characteristics used at either the individual or household level influence migrant selectivity. Finally, it tests empirically two theoretical pathways (neoclassical and network theories) relating migrant characteristics to the choice of migration destinations.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1 : Introduction

Using a three-chapter design, this dissertation addresses multiple topics of current population processes pertaining to child health, migration, and living arrangements.

The first chapter is concerned with the impact evaluation of malaria control interventions in mainland Tanzania, where malaria imposes a high disease burden on the population, especially on children. Recently, there have been several interventions that seek to control or eradicate malaria in African countries, after many experimental pilot studies established the effectiveness of delivery of insecticide-treated bednets, rapid diagnosis and treatment of malaria cases, and intermittent preventive treatment among pregnant women (Alonso et al. 1993, Binka et al. 1996, D'Alessandro et al. 1995, Greenwood et al. 1987, Ter Kuile et al. 2003). Since the early 2000's massive malaria control interventions have been implemented in Africa. At the same time, child health has witnessed tremendous improvements in most countries. Yet, the extent to which malaria control interventions have contributed to these improvements in child health is not currently known. The main reason for this knowledge gap is the difficulty in finding plausible counterfactuals in the presence of large-scale health programs that are implemented at the national level. This chapter addresses this challenge in mainland Tanzania by exploiting the timing of the dramatic scale up of interventions related to malaria control combined with the fact that pre-intervention malaria endemicity depends on ecology (as anopheles mosquitoes cannot thrive at altitudes above 1,900m (Mohammed et al. 2015, NMCP-MoHSW 2013, Taylor and Mutambu 1986)) to assess the impact of malaria control on child health. The counterfactuals are defined as areas that are naturally inhospitable for malaria mosquitoes, where malaria control activities would virtually have no impact on child health. The empirical strategy employed is a difference-in-difference approach. Using birth history data and measures of child nutritional health from the 2004-5 and 2009-10 waves of the Tanzanian Demographic and Health Survey, I estimate the specific contribution of malaria control to improvements in child mortality, morbidity, and nutritional health controlling for several socioeconomic confounders such as mother's education and mother's age.

The second chapter investigates patterns of living arrangements among the foreign-born in the United States. It is well established that living arrangements vary substantially between native-born and foreign-born populations within the United States, even net of socioeconomic and demographic differences across groups (Glick, Bean, and Van Hook 1997, Goldscheider and Bures 2003, Van Hook and Glick 2007). Immigrants' greater propensity to live in extended households is potentially a source of concern, as it may signal economic distress or a lack of incorporation into larger mainstream society (Kamo 2000, Moen and Wethington 1992). A clear understanding of the nature and source of disparate living arrangements across groups is also a critical issue in its own right, as they influence all analyses conducted at the household level; a failure to account for immigrants' tendency to live in extended households could result in an underestimation of ethno-racial inequality in a wide variety of economic outcomes, from household income to homeownership. In spite of the importance of living arrangements to stratification, the literature on the topic suffers from a number of shortcomings. First, the majority of research on the topic focuses on Latin American (particularly Mexican) and Asian immigrants. In addition, remarkable heterogeneity across immigrant groups with respect to human capital and background characteristics (Borjas 1994, Camarota 2012, Elo et al. 2015, Feliciano 2005, Jasso 2011, Jasso et al. 2004) offers new leverage in the attempt to separate economic from cultural sources of disparities in household extension (coresidence) from the native born. Another shortcoming of the household extension literature is that it often fails to distinguish between different types of extension. As Van Hook and Glick (2007) have highlighted, the determinants of vertical extension (across generations) often differ from those of horizontal extension (within generations). Finally, a common conclusion in the literature on immigrant extension is that much of the residual difference in living arrangements that remains between the foreign and native born after accounting for socio-demographic characteristics is attributable to disparate cultural preferences. However, previous studies have failed to consider how the uneven geographic distribution of groups across the country may also contribute to these disparities. Specifically, immigrant populations tend to be far more highly concentrated than the native born geographically, and are particularly over-represented in urban and coastal areas, where the cost of living is higher. It is thus important to take into consideration not only individual and household level characteristics, but also differences across groups in local context when assessing variation in extension patterns. Drawing on data from the 2001 through 2013 waves of the American Community Survey, I address these gaps in our understanding of immigrant household extension. First, I estimate up-to-date prevalence rates of extension among native-born non-Hispanic whites and multiple foreign-born groups (Mexicans, Latin Americans (excluding Mexicans), West Indians, South East Asians, Asians (excluding South East Asians), Africans, Oceanians, and Europeans), including the relative contribution of socio-demographic characteristics to household extension. Second, I assess the prevalence and predictors of extension separately for horizontal and vertical extension. Third, I investigate the extent to which variation in local context helps explain some of the residual differences across groups, net of socio-demographic differences, evident in previous studies. And finally, I examine variation in the predictors of extension across immigrant groups, separately for vertical and horizontal extension.

Finally, the third chapter examines emigrant selectivity and the factors that shape the choice of migration destinations. The rapid and sustained increase in emigration from Africa has heightened interest in the selectivity of these flows (Castagnone et al. 2015, Schoumaker, Flahaux, and Beauchemin 2015, van Dalen, Groenewold, and Schoorl 2003). Understanding the nature and scope of selective emigration out of Africa is important for both theoretical and policy reasons, since it is closely related to the ongoing debates about brain-drain and its consequences for Africa (Anarfi, Quartey, and Agyei 2010, Docquier and Rapoport 2007, 2012). The outflow of skilled emigrants is not only thought to deprive African nations from citizens who can make the most valuable contribution to the development process, but it is also thought to cause losses in public spending on education as emigrants receive education in their origin countries before migration (Anarfi, Quartey , and Agyei 2010, Docquier and Rapoport 2007, 2012). Although selective emigration out of Africa has important implications for development, it remains relatively understudied. Using data from the 2009 Ghana survey of the Migration from Africa to Europe project (MAFE), the British Labor Force Survey (waves 2010-12), and the American Community Survey (waves 2005-10), this study bridges four gaps in our understanding of selectivity among emigrants from Africa. First, this study investigates how current emigrants compare to return migrants and non-migrants in terms of their socioeconomic characteristics. Second, it explores whether emigrants to various destinations differ in their socioeconomic characteristics. Third, the study examines selectivity by using alternative measures of socioeconomic characteristics at either the individual or household level. Finally, it investigates whether Ghanaian emigrants to the United Kingdom and the United States base their migration decision on a comparison of earnings across multiple destinations or whether they rely on the presence of networks to choose their emigration destinations.

Although the three essays represent self-contained, independent articles, they share the commonality that they contribute new empirical knowledge with important policy implications for the health and wellbeing of populations in both developing and developed countries. They also utilize various methodological approaches to investigating population processes.

CHAPTER 2 : Role of Malaria Control in Improving Child Health in Mainland Tanzania: Evidence from a Rapid Policy Scale-up

Abstract

Malaria eradication is a major policy concern in developing countries, where the disease imposes a high burden on child health. Since 2004, in mainland Tanzania, malaria control interventions have experienced a rapid scale-up. Between 2004 and 2010, the budget devoted to malaria control has risen from less than $10 million to nearly $140 million. Interventions targeting malaria eradication have mainly consisted of a delivery of insecticide-treated bednets (ITN). I exploit the timing of this sharp increase in the delivery of ITN, along with the variation in malaria endemicity across the country due to ecology, to evaluate the impact of malaria control interventions on child health with an intent to treat in mainland Tanzania. My estimates suggest that, on average, malaria control interventions have helped avert approximately 17.9 deaths for every 1,000 live births. In relative terms, they have contributed to 57.7 percent of the reduction in under-five mortality between 2004 and 2010. These interventions have also improved other measures of child health, such as anemia and stunting, whose odds have been reduced by 52 and 36 percent, respectively.

2.1. Introduction

The extent to which interventions geared toward malaria control have improved child health is an important concern among policy makers and researchers, as this disease is thought to impose a high burden on children's health (Rowe et al. 2007, Smithson et al. 2015, Steketee and Campbell 2010, TMIERG 2012, WHO 2014). Many studies have investigated the impact of national malaria control in African countries, where Plasmodium falciparum, the dominant parasite strain, is responsible for high morbidity and mortality, especially among children, who typically do not have an immune system well developed against malaria (Rowe et al. 2007, Steketee and Campbell 2010). Most of these studies, relying on an after-before design combined with a "plausibility argument" (Rowe et al. 2007), have suggested possible large health benefits attributable to malaria control. For example, many have documented a downward trend in indicators such as under-five mortality and the prevalence of anemia associated with scale up malaria control interventions (Rowe et al. 2007, Smithson et al. 2015, Snow, Trape, and Marsh 2001, Steketee and Campbell 2010, TMIERG 2012).

Yet, there are three shortcomings in past studies of the impact of national malaria control on child health. First, they usually fail to provide estimates of the effect size of health improvements that can unambiguously be related to malaria control interventions (Rowe et al. 2007). In other words, they often do not quantify the specific contribution of malaria control to improvements in child health. Second, these studies do not usually control for confounders such as socioeconomic characteristics of children under study. Third, they do not investigate the contribution of malaria control to improving child nutritional health.

This paper addresses these shortcomings by using indicators of child nutritional status and by employing a difference-in-difference strategy. It exploits the timing of the dramatic scale up of interventions related to malaria control combined with the fact that malaria endemicity depends on ecology (as anopheles cannot thrive at altitude above 1,900m (Mohammed et al. 2015, NMCP-MoHSW 2013, Taylor and Mutambu 1986)) to assess the impact of malaria control on child health in mainland Tanzania. The malaria control interventions have mainly consisted of vector control activities, introduced gradually beginning in 2000. Between 2004 and 2006, a voucher scheme was implemented to provide subsidized insecticide-treated bednets to pregnant women (Hanson et al. 2008). In 2007, the voucher scheme was expanded to children under five (Hanson et al. 2008), while, in 2008, a catch-up campaign has provided more than 9 million bednets at the national level (Bonner et al. 2011). During the same time period, all-cause under-five mortality has decreased from 112 deaths per 1,000 live births in 2004 to 81 deaths in 2010 (Macro 2011, TMIERG 2012). Likewise, the prevalence of severe anemia dropped from 11.1 percent to 5.5 percent over the same time period (Macro 2011, TMIERG 2012).

Using birth history data and measures of child nutritional health from the 2004-5 and 2009-10 waves of the Tanzanian Demographic and Health Survey, I estimate the specific contribution of malaria control to improvements in health among children under five. I examine a number of indicators of child mortality, morbidity, and nutritional health and controlling for several confounders. The estimates suggest that, on average, malaria control interventions have helped avert approximately 17.9 deaths for every 1,000 live births between 2004 and 2010. In relative terms, the interventions have contributed to 57.7 percent of the reduction in under-five mortality over this time period. These interventions have also reduced morbidity and improved child nutritional health in a significant way.

The remainder of the paper is organized as follows. In Section 2.2, I present the related literature. Section 2.3 describes the context and the data. Section 2.4 presents the identification strategy, while Section 2.5 is devoted to the estimation strategy. Section 2.6 presents the results, followed by Section 2.7 that focuses on issues related to robustness checks. Section 2.8 describes threats to validity and Section 2.9 provides a cost-benefit analysis of the interventions. Finally, Section 2.10 concludes.

2.2. Background

2.2.1. Malaria Disease

The malaria disease is caused by a parasite, the Plasmodium, which attacks red cells (erythrocytes) (Clark and Cowden 2003, Centers for Disease Control 2010, Newton and Krishna 1998, World Health Organization 2000). There are five plasmodium strains in the history of human malaria around the globe, and different plasmodia tend to thrive in different geographic areas (Centers for Disease Control 2010, Hall et al. 2005). The most common strain in Africa is Plasmodium falciparum, which is transmitted by a genus of mosquitoes named anopheles. In Africa, there are seven malaria mosquito species, among which anopheles gambae, anopheles arabiensis, and anopheles funestus are the most common (Hall et al. 2005).

The transmission of Plasmodium goes from anopheles to humans and then goes back to anopheles, following a two-host dynamic (Buffet et al. 2011, Clark and Cowden 2003, Centers for Disease Control 2010, Hall et al. 2005). The transmission starts when a feeding anopheles inoculates sporozoites (motile uni-cell parasites) into a human host. After inoculation, the malaria parasites go successively through the liver and blood.There is usually a window of seven to 30 days between the infective bite and the first appearance of malaria symptoms (Centers for Disease Control 2010). These symptoms can only be detected during the blood stage when evolved forms of sporozoites (merozoites) invade erythrocytes. The transmission cycle loops back to anopheles when they take a blood meal from a human host containing sexually differentiated forms of the malaria parasites. Upon the blood ingestion, these forms of parasites undergo several stages of growth inside the anopheles, lasting between 10 to 21 days, before they reach the stage of sporozoites that may initiate another transmission cycle (Centers for Disease Control 2010).

Formal diagnosis of malaria relies on parasitemia obtained by microscopy or by rapid malaria diagnosis tests. However, residents in high malaria prevalence regions, often resort to presumptive malaria treatment in the presence of a few malaria symptoms, which come as a response of the host immune system to the invasion and destruction of red cells. Clinically, uncomplicated malaria, the most common case of malaria in Africa, is associated with fever, a sensation of cold, nausea, vomiting, weakness, and anemia. In young children, it can also cause seizures (Centers for Disease Control 2010). Uncomplicated malaria can further be associated with a feeling of general malaise, headaches, and muscular pains. On the other hand, severe malaria is associated with organ impairment and conscience impairment (Centers for Disease Control 2010, Newton and Krishna 1998, World Health Organization 2000). The most prevalent forms of severe malaria affect the brain and to a lesser extent the kidneys. The Center for Disease Control recommends that healthcare providers consider severe malaria a medical emergency, requiring a prompt comprehensive treatment (Centers for Disease Control 2010).

2.2.2. *Malaria Entomology and Ecology*

Ecological and climatic factors are important for mosquito development and survival, of these, temperature plays a crucial role (Bayoh 2001, Bayoh and Lindsay 2003, Bayoh and Lindsay 2004, Kirby and Lindsay 2004). Mosquitoes develop from eggs to adult vectors through several stages, among which the first few are aquatic. The aquatic development of mosquitoes is optimal under temperatures ranging from 22 to 28°C (Bayoh 2001, Bayoh and Lindsay 2004). At temperatures below 16 °C and above 34°C, larvae cannot produce adult mosquitoes (Bayoh 2001, Bayoh and Lindsay 2004). Mosquitoes' survivorship is the highest under temperatures between 15 to 25°C, and their survivorship declines sharply as temperature increases (Bayoh 2001). Temperature also affects the length of each developmental stage of mosquitoes. Under a tropical climate, the time span from eggs to adult vectors lasts approximately five to 14 days, while in temperate climates this length is more than 30 days[1]. Thus, the likelihood of developing the malaria disease highly depends on seasons and geographic locations.

2.2.3. *Malaria and Child Mortality*

Malaria is the third most important cause of death among children under five in Africa after pneumonia and diarrhea (Black et al. 2010). It is believed that between 16 and 60 percent of deaths occurring to those under five years of age are attributable to malaria in Africa (Alonso et al. 1993, Alonso et al. 1991, Black et al. 2010, Greenwood et al. 1987). In the last few decades, research on the role of malaria in

[1] https://www.cdc.gov/malaria/about/biology/mosquitoes/index.html

child mortality has generated a substantial literature, encompassing a variety of study designs. Several studies have relied on cross-sectional analysis of data from verbal autopsy to investigate the proportion of deaths due to malaria. For instance, using data collected with a post-mortem questionnaire (verbal autopsy), Greenwood and colleagues (1987) find that malaria accounts for 25 percent of deaths in children aged one to four (Greenwood et al. 1987). Yet, this estimate of the contribution of malaria to child mortality has not convinced everyone, as there is serious concern that it might be downwardly biased (Rowe et al. 2007, Shanks, Hay, and Bradley 2008, Snow, Trape, and Marsh 2001, Steketee and Campbell 2010). Shanks and colleagues (2008) have shown that the downward bias in the estimates of malaria-induced mortality with verbal autopsy or vital registration data can be as large as the estimates of mortality themselves, because not only does malaria increase mortality related to other infectious diseases such as pneumonia (Shanks, Hay, and Bradley 2008), it also increases the risks of mortality in children through its tendency to cause low birthweight and pre-term births (Snow, Trape, and Marsh 2001). The research on malaria-induced mortality based on observational data, grapples with the fact that the methods to capture the specific contribution of malaria to child mortality often lack both sensitivity and specificity.

The advent of insecticide-treated bednets (ITN) provides a better means to investigate the role of malaria in child mortality, and randomized controlled trials (RCT) of ITN's have been conducted in many African countries. In two studies on the impact of the use of ITN on child mortality in the rural Gambia, Alonso and colleagues (1991 and 1993) find a 60 percent decline in all-cause mortality among children aged between one and four attributable to a reduction in malaria prevalence (Alonso et al. 1993, Alonso et al. 1991). In another study in the Gambia, D'Alessandro and colleagues (1995) show that large scale delivery of ITN has led at most to a 38 percent decline in mortality among children aged one to nine.

The discrepancies in the findings across these studies are thought to stem from differences in the uptake of ITN among children under five and in malaria transmission intensity in the study regions, where transmission intensity varies from low to moderate (D'Alessandro et al. 1995). The intensity of malaria transmission is often measured by the entomological inoculation rate, referring to the mean annual or seasonal number of infective bites of malaria mosquitoes. Low to moderate transmission intensity usually refers to infective bites, ranging between one and 10 per year, while high transmission intensity is associated with more than 100 infective bites per year (Binka et al. 1996, D'Alessandro et al. 1995).

To extend these findings, many studies have sought to assess malaria-induced child mortality in areas with high malaria transmission. These studies estimate the specific contribution of malaria to child mortality as between 17 and 33 percent across various settings in Africa (Binka et al. 1996, Nevill et al. 1996, Ter Kuile et al. 2003). These studies have first demonstrated that the relative contribution of malaria control to child mortality reduction tends to be inversely proportional to transmission intensity. This reflects more the fact that mortality is relatively higher in the high transmission intensity areas compared with the low transmission areas than a substantive difference in the mortality response to malaria control. The results have also underscored the fact that cultural factors that are not supportive of a wide adoption of ITN tend to diminish the contribution of ITN delivery to reducing malaria-induced mortality. Overall, it can be inferred that despite differences in the magnitude of malaria-induced mortality from one study to the other, the RCT's have empirically established the fact that malaria is responsible for a large share of child mortality in

Africa.

Because the early evidence of a mortality reducing effect of ITN delivery is drawn from studies, providing households with free ITN, the results do not answer the other policy related question of whether ITN delivered at some price would still positively affect child health. Schellenberg and colleagues (2001) help bridge this gap in a study in rural Tanzania, where malaria transmission is high, median monthly household consumption ranges between 77 and 96 $US, and ITN are sold at $5 and treatment kits at $0.42 (Schellenberg et al. 2001). The authors report that under-five mortality has declined by 18 percent after the delivery of subsidized ITN's (Schellenberg et al. 2001). Although the study takes place in a low income area, its results are consistent with findings from prior RCT's.

The consistent findings of large health benefits from ITN delivery among children have supported scale-up malaria interventions across Africa. At the turn of the millennium, in line with the 2000 Abuja Declaration (Malaria Roll Back World Health Organization 2000), most national malaria programs have used social marketing with various schemes to promote the household use of subsidized ITN (MoHSW 2008, Rowe et al. 2007). The early program goal was malaria control. In subsequent years, some African countries have shifted to free delivery of ITN and to more ambitious interventions geared toward malaria eradication (MoHSW 2008). After a decade or so of massive interventions, the specific contribution of these efforts to improving child health is not yet clearly understood.

The few studies that aim to assess the impact of national malaria programs include research by Steketee and Campbell (2010). Reviewing data from more than 30 African countries, Steketee and Campbell (2010) find a downward sloping trend in child mortality associated with rapid scale-up of core malaria control interventions, consisting of delivery of ITN, indoor residual spraying, malaria case management, and intermittent preventive treatment of malaria during pregnancy (Steketee and Campbell 2010). They go on to suggest a plausible causal relationship between malaria control and the decrease in child mortality, whose effect size is conjectured to exceed a 20 percent reduction in all-cause under-five mortality (Steketee and Campbell 2010).

Another published investigation of the impact of malaria control on child health is the study by Smithson and colleagues (2015), implemented in mainland Tanzania, relying on an after-and-before design that uses a "plausibility argument". In this study, Smithson and colleagues (2015) assess child health with the following measures: infant and under-five mortality and prevalence of anemia (Smithson et al. 2015). The authors compare the levels of these health indicators before and after scale-up interventions and examine if the patterns of changes are consistent with biological and ecological mechanisms, whereby a decline in malaria would translate into health benefits (Smithson et al. 2015). Smithson and colleagues (2015) argue that some unspecified fraction of the decrease in under-five mortality and of the drop in the prevalence of severe anemia is attributable to malaria control (Smithson et al. 2015). In other words, the study has provided no specific estimates of the contribution of malaria control to the observed improvements in child health, which could have potentially stemmed from many confounding factors.

2.2.4. Malaria and Child Nutritional Health

In many African countries experiencing high prevalence of malaria, children's nutritional health (stunting, wasting, and low weight-for-age) is also a tremendous health challenge (De Onis, Blössner, and Borghi 2012). Poor child nutritional health is a serious public health problem with negative influences on other aspects of child health, child development, schooling, and later life outcomes (De Onis, Blössner, and Borghi 2012, Guerrant et al. 2013, Guerrant et al. 2008, Hoddinott et al. 2008). Poor nutritional status can also weaken children's immunity to infectious diseases (Eswarappa, Estrela, and Brown 2012, Guerrant et al. 2013).

The idea that malaria and child nutritional status might be related is not new (Ferreira et al. 2015, Shankar 2000). Several studies have investigated whether poor nutritional status in children is associated with frequent episodes of malaria and whether under-nourished children are exposed to a high likelihood of developing malaria (Ferreira et al. 2015, Shankar 2000). Findings from these studies suggest that malaria mortality tends to be higher among under-nourished children compared to their better fed counterparts, but that the incidence and prevalence of malaria do not seem to vary by nutritional status (Ferreira et al. 2015, Shankar 2000). Although there is no clear consensus among nutritionists about the clinical relationship between malaria and malnutrition (Arinaitwe et al. 2012, Deen, Walraven, and Von Seidlein 2002, Ehrhardt et al. 2006, Ferreira et al. 2015, Fillol et al. 2009, Friedman et al. 2003, Mitangala et al. 2013, Nyakeriga et al. 2004, Olumese et al. 1997, Renaudin 1997, Shankar 2000, Van den Broeck, Eeckels, and Vuylsteke 1993), a growing body of research has suggested that malaria control has led to improvements in child nutritional health (D'Alessandro et al. 1995, Kang et al. 2013, Snow et al. 1991, Ter Kuile et al. 2003).

In Tanzania, where the dominant parasite, Plasmodium falciparum, is particularly virulent (Buffet et al. 2011, Charlwood et al. 1997, Clark and Cowden 2003, Midega et al. 2007, Newton and Krishna 1998, Sinka et al. 2010), there are good reasons to suspect an impact of malaria control on child nutritional health. First, frequent exposure to malaria may affect child nutritional status through inadequate ingestion of nutrients. Second, malaria may affect child nutritional health through availability of resources because more household spending on health due to malaria may deplete economic resources available to all household members. Malaria may also reduce people's ability to engage in economic activities. Hence because of malaria household members of working ages may earn less. Such an adverse effect of malaria is likely to be critical for children living in poverty, whose resource constraints are binding for food consumption.

2.2.5. Contribution to the literature on impact evaluation of malaria control

Past studies seeking to assess the impact of national malaria control on child health share three shortcomings. First, they fail to provide estimates of the magnitude of gains in child health that can unambiguously be related to malaria control. Second, these studies do not often rely on multivariate analysis to control for confounders such as socioeconomic characteristics of children under study. Finally, they do not investigate the contribution of malaria control to improvements in child nutritional health.

I respond to these shortcomings by using a quasi-experimental design that relies on a difference-in-difference approach with cross-sectional data from multiple waves of the Tanzanian Demographic and Health Survey and by investigating the impact of malaria control on child nutritional health. The main contribution of this paper is to provide estimates of the specific contribution of malaria control to improvements in child mortality and nutritional health in mainland Tanzania. This novel approach to measuring the impact of malaria control is described in Section III, following the next section that presents the data and the context.

2.3. Context and Data

2.3.1. Malaria Epidemiology in mainland Tanzania

In mainland Tanzania, malaria is seasonal or perennial, with transmission intensity varying from very low to high transmission. Recent estimates indicate that 80 percent of the population lives in regions where malaria transmission is stable perennial to stable seasonal (MoHSW 2008, NMCP-MoHSW 2013). About 16 percent of the population live in areas with intermittent transmission, and the remaining four percent live in areas that are essentially malaria free (MoHSW 2008, NMCP-MoHSW 2013). Malaria incidence usually peaks during the rainy season, leading to variation in the burden of malaria across the country as some regions experience two rainy seasons, whereas others have only one (Macro 2011, Rowe et al. 2007). Plasmodium falciparum is the dominant parasite strain. The main malaria vectors are anopheles gambiae (67 percent), anopheles funestus (22 percent), and anopheles arabiensis (9 percent) (Kabula et al. 2011, Sinka et al. 2010). Anopheles arabiensis and Anopheles gambiae have the highest entomological inoculation rates. However, recent field data suggest changes in the relative distribution of malaria mosquitoes in mainland Tanzania as a result of the ongoing use of insecticides (NMCP-MoHSW 2013). As far as vectorial capacity, i.e., the proportion of anopheles recorded to be carrying Plasmodium sporozoites, is concerned, it is the highest among anopheles gambiae (25 percent) (Kabula et al. 2011).

2.3.2. Malaria Control Program in Mainland Tanzania

The malaria control interventions have benefited from large increases in funding since 2000 (MoHSW 2008). Prior to 2004, public spending on malaria control was very low; but from 2004 to 2010, it rose steadily from $6 million to nearly $140 million (TMIERG 2012). It is important to note that the pace of increase in the budget of malaria control has not been even over this time period; spending from 2007 to 2010 accounts for approximately 90 percent of the aggregate budget (TMIERG 2012). Sustained high levels of funding, prevailing since 2007, has helped implement in mainland Tanzania a three-pronged strategy geared toward malaria eradication, consisting of vector control, case management, and malaria prevention.

Vector Control

Between 2000 and 2010, vector control, mainly consisting of the delivery of long lasting insecticide-treated bednets (LLIN/ITN), has been the linchpin of malaria control in mainland Tanzania. Indeed, the delivery of these treated bednets has represented 47 percent of the $450 million devoted to malaria control (TMIERG 2012). The access of the Tanzanian population to ITN has gone up in a way that mirrors increases in the budget devoted to malaria control.

In 2004, the Government of Tanzania initiated a phase-in voucher scheme to make ITN available to pregnant women, which reached a full national coverage by 2006 (Hanson et al. 2008). The voucher scheme entails a top-up payment from beneficiary women, ranging from $0.21-0.87, in a context where the average monthly income is approximately $90 (Hanson et al. 2008). In 2007, the voucher scheme was extended to children under five. Process indicators, however, lagged behind efforts to scale up the voucher scheme. For example, the proportion of children under five, sleeping under treated nets was 26 percent in 2008 compared to a program target of 60 percen (MoHSW 2008).

In 2008, a catch-up intervention delivered 9 million ITN to children under five and pregnant women (Bonner et al. 2011). In 2010, another catch-up intervention delivered 18 million ITN to the same targets (Renggli et al. 2013). These interventions between 2008 and 2010 have caused a sharp surge in most coverage indicators related to ITN (TMIERG 2012). The use of nets jumped between 2008 and 2010 from 26 to 56 percent among children under five and from 24 to 62 percent among pregnant women (TMIERG 2012). Whereas the delivery of ITN has been implemented at the national level, there have also been other interventions, conducted locally, aimed at reducing the number of malaria mosquitoes.

Such interventions include indoor residual spraying (IRS) in two regions where large breeding sites of mosquitoes are easy to target. In 2007, the IRS started in selected districts in the regions of Kagera and Karagwe in the northwest of the country. By 2009, it was scaled up to all districts in these two regions. Other vector control interventions include environmental control and larviciding as part of the integrated malaria control program implemented in Dar Es Salam. In addition to vector control, the malaria control program has sought to improve correct diagnosis and treatment of malaria.

Malaria Case Management

Malaria case management is the second component of the malaria eradication strategy in mainland Tanzania after vector control. It includes appropriate treatment of malaria cases after correct diagnosis is provided by microscopy or rapid diagnosis tests (MoHSW 2008). A key attraction of the case management strategy is that correct and rapid treatment of malaria can not only reduce the disease burden of malaria, but can also help break the transmission cycle, contributing to reducing anopheles' vectorial capacity. Unfortunately, laboratory tests seeking to confirm malaria cases have not yet been widely available in Tanzania, and the treatment of malaria cases often relies on clinical diagnosis (MoHSW 2008, TMIERG 2012). Malaria cases are managed according to the new guidelines of the World Health Organization, recommending the use of artemisinin combination (ACT) as the first-line therapy for malaria. Quinine is recommended for severe forms of malaria. Since 2009, ACT drugs have been mostly available at more or less affordable prices (if not subsided), despite a stock-out in 2010.

Intermittent Preventive Treatment

The third component of the strategy to eradicate malaria in mainland Tanzania consists of malaria prevention. Preventive interventions mainly focus on reducing malaria cases in pregnant women, as malaria during pregnancy can affect women's health and birth outcomes. Malaria prevention in pregnant women has had a historically low uptake in mainland Tanzania (MoHSW 2008, TMIERG 2012). In 2010, 26 percent of the pregnant women went through the complete course of prevention, consisting of two therapeutic doses (MoHSW 2008, TMIERG 2012). This represents a modest improvement of 5 percentage points since 2004, and still falls short of the national target of 80 percent (MoHSW 2008, TMIERG 2012). The conflicting directives from the antenatal health system and the malaria control program may help explain why most women only take the first therapeutic dose, as there tends to be some confusion about the gestational term at which the first dose can be administered and the correct time spacing between the two doses (TMIERG 2012).

2.3.3. Data

I investigate the role of malaria control in improving child health in mainland Tanzania, using micro-level data from two waves of the Demographic and Health Survey (DHS), fielded respectively from October 2004 to February 2005 and from December 2009 to May 2010 (Macro 2005, 2011). These data are collected by the Tanzanian National Bureau of Statistics with technical assistance from Macro International Inc.; they are gathered from two independent, nationally representative samples of 9,852 and 9,741 households respectively. In each survey wave, the response rates are above 99 percent at the household level and 96 percent at the level of the individual women, with slightly higher rates (in an order of magnitude of 0.2) in rural areas compared to urban ones (Macro 2005, 2011). In most areas of the country, the time period of the data collection overlaps with the rainy season (when malaria prevalence is high), which helps measure malaria and its effects on child health. The DHS survey collects birth histories of all women of reproductive ages (15-49 years), including an indicator of whether the children born to a given mother are alive at the time of the survey. It also provides information on the child anthropometric measures and a wealth of household background characteristics.

I merge district-level data on population-adjusted malaria prevalence among children aged two to ten in 2000 (retrieved from the report titled “An Epidemiological Profile of Malaria and its Control in Mainland Tanzania” (NMCP-MoHSW 2013)) to the DHS data. This prevalence measure has been used in many settings as a measure of the burden of the malaria disease at the population level (NMCP-MoHSW

2013). It is obtained from a Bayesian hierarchical regression, using data from 2,193 time-surveys at 1,447 distinct locations in Tanzania (NMCP-MoHSW 2013). Data on malaria prevalence have provided a means to explore likely health responses to malaria control.

I restrict the DHS sample to the subset of children under five, who have no missing information on the explanatory variables (presented below). In total, 5.3 percent of children under-five were excluded. Throughout the analysis, the sample varies in size across health outcomes because some of these outcomes have missing information. Table 2.1 presents the analysis sample by survival status and year of birth. Data on child mortality have no missing information, yielding a pooled sample consisting of 12,612 children under five, which includes 11,677 surviving children. This sample is slightly unbalanced between the two waves as it consists of 52.2 and 47.8 percent in the 2004 and the 2010 DHS, respectively. The pooled sample pertains to children, mostly born between 2000-2004 in the 2004 DHS (only 0.7 and 0.4 percent were born respectively in 1999 and 2005), and to children born between 2005 and 2010 in the 2010 DHS.

For the other measures of child health, the analysis sample is restricted to children who are alive at the time of the survey. The anemia variable has no missing data, resulting in a sample of 10,463 children between six and 59 months of ages, there are missing values in child weight and height. Missing information affects 10.6 percent of children alive at these ages, yielding a sample of 9,706 children under five for the analysis of child nutritional health. For robustness checks, I supplement the analysis with analogous data from the 1999 DHS, consisting of 2,406 children under five, who are either alive or dead for the analysis pertaining to the under-five mortality. The rationale of this step in the analysis is to check whether the program effects are absent between 1999 and 2004, as no program has been implement during this time period.

2.3.4. Dependent Variables

I examine the impact of the malaria control program on four measures of child health. First, I investigate the impact of malaria control on under-five mortality, which is a key impact measure recommended by the Roll Back Malaria monitoring and evaluation board (Rowe et al. 2007, TMIERG 2012). I estimate under-five mortality with the pooled DHS data among children under five. (section 2.5 on the Estimation Strategy describes the approach to modeling under-five mortality). Survival information is provided by the observation time recorded in months and a child's survival status, indicating whether a child is alive at the time of the survey and the month of death for those deceased. Second, I examine the impact of malaria control on the prevalence of moderate anemia. The level of anemia is closely related to malaria, which destroys blood red cells and tends to cause anemia in children (Buffet et al. 2011, Clark and Cowden 2003, Newton and Krishna 1998). Moderate anemia is defined as an altitude-adjusted level of hemoglobin between 7.0-9.9g/dL. This variable is collected among children aged 6 to 59 months, who are alive at the time of the interview.

I further investigate the impact of malaria control on stunting and low weight-for-age because these measures of child nutritional health, especially stunting, are indicative of children's exposure to infectious diseases (Eswarappa, Estrela, and Brown 2012, Guerrant et al. 2013). I estimate the prevalence of stunting and low weight-for-age, using the World Health Organization growth charts and recommended cutoffs, together with anthropometric data from the DHS. Weight and height are available for the children aged 6 to 59 months, who are alive at the time of the interview.

Following the WHO guidelines, I consider children falling under two standard deviations of their standardized height-for-age or weight-for-age to define stunting and low weight-for-age, respectively (Macro 2011).

2.3.5. Explanatory variables

Key explanatory variables are related to exposure to malaria dummy pre and post-intervention periods. These variables are described in section 2.4, which explains the identification strategy. In addition, I use many proximate determinants of child health identified in previous research (Alexandre et al. 2015, Behrman and Deolalikar 1988, Berger, Fahrmeir, and Klasen 2002, Rosenzweig and Schultz 1983, Rosenzweig and Wolpin 1988, Strauss and Thomas 1998, Trussell and Hammerslough 1983). However, because I am interested in evaluating the impact of malaria control, the choice of control variables is guided by the requirement that these variables may not be directly altered by the malaria control interventions. Thus, the exogenous control variables include child's sex; child's age (recorded in months for the analysis dealing with anemia and nutritional health); mother's education (three binary indicators of no education, primary, secondary or higher); education of mother's partner (a proxy for father's education, measured by the same three binary indicators as for mother's education); household head's sex (male=1); an indicator of whether child lives in an urban area; and five indicators of living standards, capturing whether a household has access to clean water, whether a household has access to improved sanitation, whether the wall, the floor or the roof of the dwelling place is made from modern materials, respectively. I also use season of birth (rainy season=1), which is constructed from the month of birth, place of birth, and information on the timing of the dry and rainy seasons in mainland Tanzania. Note that place of birth is inferred from the mother's place of residence at the time of the survey, as there is no other information in the DHS dataset to indicate a child's place of birth. I further use an indicator of whether the mother attended antenatal care during pregnancy and an indicator of whether the mother reported distance as a big concern for the use of healthcare for herself. These variables aim to capture information about access to healthcare or alternatively the price factors associated with child health. I use mother's age, captured by four binary indicators, to control for the risks related to giving birth at extreme ages (very young or older ages).

To describe the changes in program uptake between survey waves, two variables about program coverage are presented. These variables include the proportion of children sleeping under an insecticide-treated bednet the night before the survey and a variable at the household level, indicating for each child whether all children in her/his household sleep under a treated bednet. Because these program coverage variables are process indicators, they are not used in the multivariate analysis of the impact of malaria.

2.4. Identification Strategy

The evaluation of large scale health programs is frequently complicated by the fact that program assignment may not be randomized. To address this difficulty, researchers have often used plausibly exogenous criteria to define the treatment and control groups with observational data. In this study, the Plasmodium falciparum parasitemia at the population level, which helps characterize malaria endemicity[2], and the survey years provide the basis for using a difference-in-difference approach

[2]There is a close correspondence between malaria endemicity and the intensity of malaria transmission.

with an intent to treat to identify the impact of malaria control on child health in mainland Tanzania. I exploit the fact that low malaria endemicity, prevailing prior to any vector control interventions, in a given area reflects the fact that this area is inhospitable for malaria vectors. In Tanzania, highlands usually experience very low malaria endemicity, reflecting the inability of anopheles there to survive or to contribute to the vector-host transmission, as they are exposed to low average temperatures (Mohammed et al. 2015, NMCP-MoHSW 2013, Taylor and Mutambu 1986).

In this paper, the intensity of malaria endemicity, before any anthropic actions, seeking to modify the epidemiological profile of malaria, is measured by the 2000 population-adjusted prevalence of malaria parasites in children aged two to ten[3] measured at the district level. Figure 2.1 illustrates malaria endemicity in mainland Tanzania based on this measure. I define two groups of areas that experience two distinct regimes of exposure to the malaria disease by using a cutoff of five percent of the population-adjusted prevalence of Plasmodium falciparum. The choice of this point value of five percent is motivated by the fact that below this cutoff, malaria is either absent or hypo-endemic (NMCP-MoHSW 2013). Nine percent of the children in the analysis sample live in the low endemicity areas, where the population adjusted prevalence is below five percent; while the rest of the sample lives in the high endemicity areas, where the population adjusted prevalence vary from five to 93 percent. Although the range of prevalence in the high endemicity areas is wide, it reflects thresholds at which malaria can be clinically damaging at the population level (NMCP-MoHSW 2013).

The low endemicity areas are then considered control areas, whereas the high

[3]The measure is highly negatively correlated with altitude (Pearson correlation = -0.567; p-value =0.000).

endemicity areas are the intervention areas. The rationale for this choice is that the impact of malaria control will be minimal in the areas with very low exposure to malaria, as cases of malaria are unlikely with malaria being absent or very rare (sleeping under a bednet in areas with no anopheles, would have no protective role against malaria). However, depending on the intensity of migration between the high and low endemicity areas, correct and rapid diagnosis could make a difference in a child's health related to malaria. In any case, net in-migration to the areas of low malaria endemicity would lead to underestimate of the impact of malaria control on child health. One important thing to note is that the "treatment" is not defined as being exposed to malaria control interventions. It is rather defined as living in areas, where after being exposed to the malaria control interventions, households would be expected to improve their children's health through the interventions.

The validity of this approach to assess the impact of malaria control rests on the assumption that there would be no change in the epidemiology of malaria in mainland Tanzania during the time span between 2004 and 2010 if there were no malaria control. Put differently, it must be the case that natural conditions did not change independently in such a way that they would have modified malaria epidemiology over this time period. Indeed, between 2000 and 2010, the ecological profile of Tanzania has not changed in a way that would have affected the environmental suitability of anopheles in Tanzania (NMCP-MoHSW 2013).

The timing of the successive waves of the DHS maps almost perfectly onto the before and after intervention time period. I consider children under five, in the 2004 DHS to be in the pre-intervention period, while children under five in the 2010 DHS are considered to be in the post-intervention period. The timing of the 2010 DHS is fortunate in the sense that children have been exposed to the risk of malaria over a few rainy seasons (periods of high transmission intensity) have passed. This allocation rule, defining the before and after intervention, might be seen as suboptimal because some children, born between 2004 and 2005, have benefited from some interventions of the malaria control program, mainly associated with the voucher scheme (Hanson et al. 2008). However, less than 0.5 percent of the children in the 2004 DHS were born in 2005. Furthermore, coverage indicators of the malaria control program before 2006, show that most program interventions have had low coverage at the national level (MoHSW 2003, 2008, TMIERG 2012). Remember that large scale malaria control interventions have been implemented from 2006 onward.

With these definitions of intervention and control areas, and of pre and post-intervention periods, one is equipped to use a difference-in-difference (DD) estimator. Equation (2.1) below illustrates the DD approach:

$$Y_i = \beta_0 + \beta_1 (M_i * T_i) + \beta_2 M_i + \beta_3 T_i + \varepsilon_i \tag{2.1}$$

where:

1. Y_i is a measure of child health;
2. M_i =1(Child lives in a high malaria endemicity area);
3. T_i = 1(Survey year is 2009-10).

The parameter β_1 is of a prime importance in this analysis as it identifies the impact of post-intervention exposure relative to exposure before the national-level program was rolled out. The identification strategy using a DD estimator accounts for the fact that the indicators of child health of interest are not malaria specific. What the DD estimator is picking up is the specific contribution of malaria control to reduction in mortality. It allows for differences in the baseline in each endemicity group. However, the DD estimator does not allow for differences in the slope of the trends in indicators, stemming from factors other than those defined by the treatment (Angrist and Pischke 2008, Shadish, Cook, and Campbell 2002, Todd 2007). In Section 2.9, I shall explore the validity of this assumption. Hence, the current identification strategy helps quantify the causal of the impact of malaria control on child health. Put differently, at the population level, the coefficient β_1, in Equation (2.1), provides a causal measure of malaria control interventions on child health (Y). In the next section, I present ways to estimate the effect of the malaria program on child health.

2.5. Estimation Strategy

I use a continuous-time duration model[4] to investigate under-five mortality (Allison 2014, Amemiya 1985, Cleves 2010, Van den Berg 2001). In the DHS dataset, the observation time is recorded in a discrete form (months), which implies possible ties. Such ties are handled with the Efron approximation procedures (Borucka 2014, Efron 1977, Hsieh 1995). Methods for duration analysis are particularly well-suited for analyzing mortality data mainly because they account for censoring, arising from the fact that not all children are observed for the same time span (Allison 2014, Amemiya 1985, Cleves 2010, Klein et al. 2016, Therneau and Grambsch 2000, Van den Berg

[4]In section 2.7, I run a robustness analysis with a discrete-time duration model.

2001). As Trussell and Hammerslough (1983, p.2) have pointed out, the duration analysis methods provide a means to investigate child mortality in a way that is consistent with the life table approach (Trussell and Hammerslough 1983).

The risks of mortality are estimated with survival information as the input. For a given child i, survival information is captured by the pair (t_i,d_i), where t_i is an integer between 0 and 59 (inclusive), representing a child observation time recorded in months, and d_i is an indicator of whether child i is alive at the date of the interview. Hence, for a child, who is alive, d_i=0 and t_i is her current age, while for a deceased child d_i=1 and t_i is her/his age at death. Children who are still alive at the time of the interview are censored. Censoring is random as the observation time for each child is a random variable, depending on her date of birth and the date of the interview. I assume that censoring is non-informative. This is a plausible assumption because the date of the interview cannot be systematically related to the survival status of every child.

There are several alternatives for modelling event time. I utilize a Cox proportional hazard model to model under-five mortality because it allows for an arbitrary form in the baseline hazard function. The Cox proportional hazard model also helps focus on the hazard functions, which are other ways of expressing the death rates commonly used in the demographic literature.

I define the hazard of death among children under-five as a function of their socioeconomic characteristics in the following way:

$$\ln(h(t_i|Z_i)) = \ln(h_0(t_i)) + \beta_1 M_i * T_i + \beta_2 M_i + \beta_3 T_i + \beta_4 X_i \qquad (2.2)$$

Where:

1. $h_0(.)$ is the arbitrary baseline hazard function;

2. t_i is the observation time of child i;

3. $Z_i=(M_i, T_i, X_i)$, with:

   (a) M_i =1(Child lives in a high malaria endemicity area);

   (b) T_i = 1(Survey year is 2009-10);

   (c) X_i the vector of control variables.

Given the fact that:

$$h(t_i|Z_i)) = -\frac{S'(t_i|Z_i))}{S(t_i|Z_i))} \tag{2.3}$$

Where: and $S\,(t_i \mid Z_i) = 1 - F(t_i)$ with $F(t_i)$ the cumulative distribution function of t and $S'\,(t) = \frac{dS(t)}{dt}$

In the Cox model, the survival function, S(.), is a function of the parameters in equation (2.3) and of the baseline survival probability $S_0(.)$, all of which can be estimated from the data. When there are ties, denote π_i the conditional probability, solution to the following equation:

$$\sum_{k\ \in\ \Delta_i} \frac{e^{Z_k\beta}}{1-\pi_i{}^{e^{Z_k\beta}}} = \sum_{k\ \in\ \mathrm{R}_i} e^{Z_k\beta}$$

Where:

1. Δ_i is the set of the cases that failed at t_i;

2. β is the stack vector of parameters;

3. R_i is the risk set at time t_i.

$$S_0\,(t \mid Z) = \prod_{i:t_{(i)}<t} \pi_i \tag{2.4}$$

The survival probability is then defined as:

$$S_i\,(t \mid Z) = \left[S_0(t)\right]^{e^{Z_i \beta}} \tag{2.5}$$

Although the Cox proportional hazard model is very flexible, it imposes proportionality at all points in time in the hazards at all levels of the covariates. This restrictive assumption, which may be violated in practice, is usually tested with a test based on the Shoenfeld residuals (Allison 2014, Cleves 2010). Table 2.2 presents the results of this test with the full model, controlling for all covariates. The test of the proportional hazard assumption leads to a wide rejection of the proportionality assumption. Because the outcome of the Shoenfeld residuals test is sensitive to the sample size, it is also important to examine the magnitude of the correlations between these residuals and each covariate and to examine the number of covariates implicated in the association before making definite conclusions about the consequences of the proportionality assumption. The results from the table indicate that these correlations are relatively small and that only three covariates (the indicator of whether a child's mother attends antenatal care during child's pregnancy, whether a household is male headed, and whether a mother reports distance a big concern for her use of healthcare) have a statistically significant association with the residuals, suggesting that the implications of the violation of the proportionality assumption are likely to induce little change in the estimated parameters.

However, I relax this assumption by including new predictors that control for the interaction between time or a function thereof and the covariates with a statistically significant association with the Shoenfeld residuals. Specifically, I proceed with a Cox model that controls for the interaction between the logarithm of time t_i and each of the three covariates, causing a statistically significant association with the Shoenfeld residuals in the following fashion:

$$\ln(h(t_i|Z'_i)) = \ln(h_0\,(t_i)) + \beta_1 M_i * T_i + \beta_2 M_i + \beta_3 T_i + \beta_4 X_i + \beta_5 ln(t_i) * Z_{2i} \quad (2.6)$$

Where:

1. $h_0(.)$ is the arbitrary baseline hazard function;
2. t_i is the observation time of child i;
3. Z'_i=(M_i, T_i, X_i, Z_{2i}), with:
   (a) M_i =1(Child lives in a high malaria endemicity area);
   (b) T_i = 1(Survey year is 2009-10);
   (c) X_i the vector of control variables.
   (d) With Z_{2i} being the vector of an indicator of whether a mother attends antenatal care during a child's pregnancy, whether a household is male headed, and whether a mother reports distance as a big concern for her use of healthcare.

Another requirement of a causal analysis with the Cox regression is that there be no statistically significant interaction between the treatment variable and the covariates (Klein et al. 2016). The data meet the requirement of no statistically significant interactions between the treatment variable and the socioeconomic control variables (*results not shown*).

As far as anemia, stunting, and low weight-for-age are concerned, a logistic regression is used to estimate the effect of the malaria control interventions on their prevalence, as these measures are binary variables. Equation (2.7) below presents the specification of the logistic regression with the full set of covariates.

$$Y_i = \beta_0 + \beta_1 (M_i * T_i) + \beta_2 M_i + \beta_3 T_i + \beta_4 X_i + \varepsilon_i \tag{2.7}$$

where:

1. Y_i is a binary measure of child health;
2. M_i =1(Child lives in a high malaria endemicity area);
3. $T_i = 1$(Survey year is 2009-10);
4. X_i the vector of control variables;
5. and ε_i is an idiosyncratic disturbance term following the logistic distribution.

For all measures of child health, I begin by estimating a specification with no control variables, as shown in Equation (2.1). Then, for under-five mortality, I estimate the specification with the full set of the control variables, illustrated in Equations (2.2) and (2.6). To recover estimates of hazard functions and the survival probabilities, Equation (2.3) and (2.5) are used respectively. The program effect on survival is computed by differencing the mean survival probability between the two levels of the interaction between the indicator of high malaria endemicity and the indicator of year 2010. Finally, the specifications in Equations (2.1) and (2.7) are used to estimate the odds of anemia, stunting, and low weight-for-age. The next section presents the bivariate relationship between the outcomes and malaria prevalence, sample characteristics, description of survivorship, and estimation results.

2.6. Results

2.6.1. Malaria Prevalence and Child Health

Table 2.3 shows unadjusted correlations between measures of child health and population-adjusted falciparum prevalence (PFPR) in children aged 2 to 10 years. The results suggest a nonlinear and positive relationship between malaria prevalence and poor child health. Panel A presents the results pertaining to child survival. In 2004, i.e. before large scale malaria interventions, as shown in Column (1) of Panel A, there is a statistically insignificant positive relationship between child mortality and the prevalence of malaria. The absence of statistical evidence of this relationship hints at a possible nonlinear relationship between under-five mortality and these levels of malaria prevalence. In Column (2) a nonlinear relationship is explored with four binary variables indicating various levels of malaria prevalence, which were chosen in reference to epidemiological data that define the clinical experience of malaria at the population level in Tanzania. The estimates suggest a nonlinear, statistically significant relationship between PFPR and under-five mortality in 2004, where high levels of mortality are associated with high levels of malaria prevalence, with clear increasing gradient up to 50 percent. The last two columns of the table present the same set of estimates for the year 2010. Contrary to 2004, the continuous correlation between PFPR and child mortality is statistically significant, suggesting a weak positive relationship between the malaria prevalence and under-five mortality. Yet, the nonlinear relationship does not support the same interpretation as areas with low levels of malaria prevalence experience lower levels of under-five mortality than areas of very low prevalence. On the other hand, there is no evidence of difference in the level of mortality between areas experiencing high malaria prevalence and very low prevalence. Prima facie, these results provide a strong support to the idea of a positive effect of malaria control on under-five mortality, since they suggest a sharp mortality decline in areas previously experiencing high levels of mortality before the interventions, which are not expected to affect under-five mortality in the low endemicity (very low prevalence) areas.

Panel B of Table 2.3 presents unadjusted coefficients of a linear regression of levels of the density of hemoglobin (anemia) on the PFPR. As shown, in both 2004 and 2010, there is a negative relationship between these two variables, highlighting the clinical fact that malaria tends to cause anemia in children. The strength of the association decreases between 2004 and 2010, suggesting some improvements attribuatable to the malaria program. Similarly, the results from the binary indicators of malaria prevalence show a highly nonlinear relationship, whereby high levels of malaria prevalence are associated with low density of hemoglobin in children.

Panels C and D show the unadjusted coefficients with z-scores of height-for-age and weight-for-age for years 2004 and 2010. The results indicate evidence of a non-linear relationship between PFPR and the z-scores of height-for-age and weight-for-age, respectively. The estimates on the binary variables indicating cutoffs in the levels of malaria prevalence display a negative relationship between the PFPR and these measures of child nutritional health. Interestingly, such a relationship is not clearly seen in the unadjusted coefficients using a continuous measure of the PFPR. The changes in the malaria prevalence betwwen 2004 and 2010 cannot be directly inferred from these estimates.

2.6.2. Sample Description

Table 2.4 presents the sample characteristics by survey year and malaria endemicity. Most characteristics vary across the four analysis groups. Panel A presents indicators of program coverage. As seen in the table, program coverage has tremendously increased between 2004 and 2010. In 2004, the proportion of the children living in a household with a treated bednet is 0.15 and 0.13 in the low and high endemicity areas, respectively. In 2010, this proportion is 0.56 and 0.67 in the low and high endemicity areas, respectively. In 2004, the proportion of the children living in a household with bednets where all the children sleep under a treated bednet is 0.22 in the low endemicity areas compared with 0.23 in the high endemicity areas. In 2010, this proportion reaches 0.50 in and 0.63 in the low and high endemicity areas, respectively. The results presented above about program coverage illustrate the dramatic scale-up in the interventions related to malaria control with possible large health benefits for children.

Panel B shows a summary of the characteristics related to anemia and the children's nutritional health. The prevalence of moderate anemia has decreased significantly between 2004 and 2010 in the high endemicity areas. In 2004, this prevalence is 0.26 and 0.45 in the low and high endemicity areas, respectively; while in 2010, the prevalence of anemia in the low and high endemicity areas is 0.24 and 0.28, respectively. The results for children's anthropometric measures indicate some improvements in child nutritional health between 2004 and 2010. In 2010, the mean z-score for children's weight in the low and high endemicity areas is 0.08 and 0.02, respectively, whereas in 2004, it is -0.03 and -0.02 in the low and high endemicity areas, respectively. Similarly, children's height has improved between the two survey waves. In 2004, the average z-score of children's height is 0.00 (the sample mean) in the low endemicity areas compared with -0.04 in the high malaria areas. In 2010, this mean z-score of children's height is 0.06 in the low endemicity areas compared with 0.04 in the high endemicity areas. In terms of nutritional health, in 2004, the prevalence of stunting (height-for-age <2 SD) is 0.42 in the low endemicity areas and 0.50 in the high endemicity areas. In 2010, this prevalence is 0.49 and 0.46 in the low and high endemicity areas, respectively. While in 2004, the proportion of the children with a low weight-for-age is 0.15 and 0.18 in the low and high endemicity areas, respectively, in 2010, this proportion is 0.20 and 0.17, respectively in the low and high endemicity areas. The descriptive results on the children's anthropometric measures show that child nutritional health has improved in the high endemicity areas, whereas it has deteriorated in the low endemicity areas.

Panel C presents the characteristics of the children and of their parents. It reveals some imbalance in the levels of many indicators and correlates of child health by survey year and endemicity. In 2004, while the mean age of the children is 28.6 months and 29.2 months in the low and high endemicity areas, respectively, in 2010, this mean children's age is 30.2 months and 29.2 months in the low and high endemicity areas, respectively. In both survey waves, the proportion of the children born in the rainy season is higher in the high malaria endemicity areas than in the low malaria endemicity areas. In 2004 for example, this proportion is 0.61 in the high endemicity areas and 0.55 in the low endemicity areas. In 2004, the mothers in the low endemicity areas are, on average, younger than those in the high endemicity areas, while in

2010, the opposite is true. In 2004 and 2010, the proportion of the mothers who have at least some educational attainment is higher in the low endemicity areas than in the high endemicity areas. In both survey waves, there are no notable differences between the endemicity areas in the proportion of the male headed households. In each survey wave, the relative proportion of children living in the urban areas is higher in the low endemicity areas compared with the high endemicity areas. In both years, there are differences in the living conditions of the children by endemicity. In 2004, while the proportion of the children in the low endemicity areas who live in households that have access to clean drinking water is 0.57, this proportion for those living in such households in the high endemicity areas is 0.46. In 2010, the relative proportion of children living in households with access to clean water is 0.67 in the high endemicity compared with 0.60 in the low endemicity areas. The quality of the main material of the houses has improved between 2004 and 2010 in the low and high endemicity areas. In each year, the low endemicity areas are relatively more advantaged than the high endemicity areas in terms of socioeconomic status.

2.6.3. Child Survivorship

Figure 2.2 illustrates the Kaplan-Meier survival probabilities of the four main analysis groups, consisting of children under five living in the low malaria endemicity areas in 2004 and 2010, and those living in the high malaria endemicity areas in 2004 and 2010. Survivorship is higher among the children under five in the areas of low malaria endemicity (2010 and 2004) compared with the children in the high malaria endemicity areas. Between 2004 and 2010, there have been tremendous survival gains in the high endemicity areas. In 2010, the probability of surviving from zero to 60 months among children in the high endemicity areas is 0.920 compared to 0.879 in 2004, while this probability decreases slightly from 0.939 in 2010 to 0.932 in

2004, in the low malaria endemicity areas. Several log-rank tests uncover statistically significant between-group differences. While survivorship has improved between 2004 and 2010 in the high malaria endemicity areas (p-value=0.000), there is no statistical evidence for any change between 2004 and 2010 in the low endemicity areas (p-value=0.212). There are statistically significant differences in survivorship between survey waves (p-value=0.000), showing an improvement in survival probability of children from 2010 to 2004. Finally, a log-rank test indicates no statistically significant difference in survivorship between the low and high endemicity areas in 2010 (p-value=0.835). These bivariate results suggest that survival gains between 2004 and 2010 are mostly the results of rapid improvements in child mortality in the high endemicity areas, hinting at a likely positive impact of the malaria program. The Cox regression of child mortality closely investigates this issue.

2.6.4. Estimation Results

Table 2.5 presents hazard ratios from a series of Cox regressions of under-five mortality controlling for all explanatory variables. Column (1) reports estimates of the base model with no covariates. The coefficient on the indicator of the interaction between high endemicity and survey year 2010 is statistically significant at the 5 percent level. Remember that this is the main coefficient of interest as it indicates the impact of malaria control on under-five mortality. Column (2) provides estimates of the full model. The coefficient on the indicator of the above mentioned interaction is highly statistically significant, showing that there is some imbalance in the level of key proximate determinants of child mortality among the four analysis groups. This column indicates that malaria control interventions have reduced the risks of mortality among children under five by 51 percent holding constant all other covariates. Put differently, for every 1,000 live births, malaria control interventions have helped avert 17.9 deaths on average. This is a large effect after controlling for the set of controls. Although the coefficient referring to the program effect changes little between Columns (1) and (2), there is qualitatively an important difference between these two columns. The estimate becomes more precise, resulting in highly statistically significant effect. The results also show that net of the malaria interventions and the effect of the set of control variables, there is no difference between the pre and post-intervention periods in child mortality, while children in high malaria endemicity areas still face risks of mortality as high as 90 percent compared to their counterparts in the low malaria endemicity areas. Column (4) of the table presents results from the full model that allows for an interaction between the analysis time and the indicators of whether a child's mother attends antenatal care during the child's pregnancy, whether a household is male headed, and whether a mother reports distance as a big concern for her use of healthcare. Although these interaction terms are statistically significant, they do not affect the estimates of the impact of malaria control found in Column (2). It is worth stressing again that there is no evidence of age variation in the effect of the estimates of the program impact (Table 2).2.

Figure 2.3 illustrates the estimated survivor for the treatment and counterfactual groups. In the figure, "High endemicity*1(2010)" represents the variable indicating whether the children live in the areas of high malaria endemicity in 2010. The figure highlights the fact that the interventions have tremendously reduced the risks of death at all ages among children living in the high malaria endemicity areas in 2010 compared to the rest of children, holding all other covariates at their means. I next turn to the results from the logistic regression of the other measures of child health.

Table 2.6 presents the results of a series of logistic regressions of the indicators of moderate anemia and child nutritional health. The patterns in the results are similar to what has been observed for under-five mortality; i.e. the effects of malaria control interventions are statistically significant. The full models with a binary measure of treatment provide an assessment of the malaria control impact. Column (4) presents the adjusted odds ratios of anemia resulting from the malaria control interventions. The column indicates a reduction in the odds of moderate anemia by 52 percent attributable to the malaria control interventions. While the odds of anemia are more than twice as high in the high malaria endemicity areas as in the low malaria endemicity areas. Column (5) shows that malaria control interventions have reduced the odds of stunting by 36 percent. Column (6) presents the results pertaining to low weight-for-age. It indicates that the malaria control interventions have caused a 33 percent decline in the odds of children having low weight-for-age. Overall, the results from this section suggest a large positive impact of the malaria control interventions on multiple indicators of child health.

2.7. Robustness Checks

Although the identification strategy leaves little room for omitted factors to bias the results, there are four main concerns with the estimation strategy that might affect the results, especially for under-five mortality. First, a continuous-time modeling of under-five mortality may not be appropriate because of the existence of ties, which tend to attenuate the estimates toward zero (Hsieh 1995); but as the highest proportion of tied observations is 3.8 percent, these ties are unlikely to significantly affect the results. To evaluate this conjecture, I use alternatively a discrete-time modelling of under-five mortality with a logistic regression, controlling for child age, to explore the impact of malaria control interventions on under-five mortality. Table 2.7 presents the results from a series of regressions dealing with robustness checks. Column (1) of the table presents the results of a logistic regression of child survival status on the full set of covariates, including time (child age). They suggest that the malaria control interventions have reduced the odds of child mortality by 52 percent, which is essentially the results from the Cox regression presented in Table 2.5. Second, previous research on child health has highlighted the protective role of breastfeeding against mortality; but concerns about the endogeneity of this variable with malaria have led to the exclusion of breastfeeding from the models presented above. This omission may have biased the estimates (Elo and Miller 1991). As a robustness check, I include breastfeeding into the model presented in Equation (2.3). The inclusion of breastfeeding in the model requires some additional care because of its dynamic nature. In the DHS data, the duration of breastfeeding is collected for all children irrespective of whether they are alive or deceased at the date of the interview. For children who are breastfed at the time of the interview, the value of this variable is set to be the child's age. Because the length of breastfeeding depends on child's age and cannot exceed a child's current age, I use a time-varying binary variable that takes on the value one during the months when a child is breastfed and zero otherwise (Berger, Fahrmeir, and Klasen 2002, Trussell and Hammerslough 1983). The episode splitting approach is used to set up the data in a way that creates multiple records per child, splitting each child's observation time into a set of one-month-long intervals (Cleves 2010, Allison 2014). Furthermore, because there is missing information in the breastfeeding variable, to provide a valid comparison with the model including this variable, I need to re-estimate the Equation (2.3) with a sample, excluding cases with missing data on breastfeed. Column (2) of Table 2.7 provides estimates for such a model, while in column (3) of Table 2.7, I present the results from a full model that includes breastfeeding. The column indicates that the contribution of the malaria program to reducing the risks of mortality are almost identical to those reported in the preceding column.

Third, the models estimated for under-five mortality do not control for all possible factors that may affect child mortality, leaving room for unobserved heterogeneity. Unobserved heterogeneity can cause an artificial decline in the shape of the hazard function over time (Vaupel, Manton, and Stallard 1979). It may also cause the so-called heterogeneity shrinkage in the estimates. I account for unobserved heterogeneity with a fixed-effects model (Allison 1996) at the ward level, representing the administrative unit below the district level. Column (4) of Table 2.7 presents the results of the fixed-effects model. These results indicate that the contribution of the program to reducing child mortality is almost the same as the estimates found in Table 2.5, as there is a 48 percent reduction in the risk of mortality among children under-five attributable to the malaria control interventions. Overall, these checks underscore the robustness of the results to changes in the assumptions about the estimation procedures.

Yet, finally, there still remains concern about the appropriateness of the definition of the intervention groups. Put differently, malaria endemicity may not have adequately captured the exposure to the malaria control interventions. This is unlikely for two reasons. First, the consistency of the estimated effects across measures of child health provides a first check of plausibility that malaria endemicity has adequately captured the likely impact of the program on child health. The estimated effects are consistent with the mechanisms, whereby malaria can affect child health because the results uncover a close connection between the decline in anemia and reduction in malaria-specific mortality. Second, I check the assumption that there should be no effects of malaria control on any measure of child health between 1999 and 2004, because malaria interventions have not yet been implemented at a scale that would have produced a sizable impact at the national level (Smithson et al. 2015, TMIERG 2012). Table 2.8 presents results from a placebo regression with analogous data from the 1999 and 2004 DHS (there are fewer covariates in the table compared to the full model because the data about some covariates were not collected in 1999). The table shows no impact of malaria control on measures of child health between 1999 and 2004, as it should be.

2.8. Threats to validity

There are two sets of possible confounding factors for the estimated effects. One possible concern with attributing the estimated effects to malaria control is the fact that there may have been some reduction in the virulence of malaria-related morbidity and mortality due to ecological reasons. Yet, data on malaria epidemiology fail to provide any evidence of waning virulence of malaria (NMCP-MoHSW 2013, Steketee and Campbell 2010). On the contrary, evidence suggests increased resistance of malaria vectors to some insecticides used to treat bednets (NMCP-MoHSW 2013). Another concern about the internal validity of these results has to do with the fact that factors other than malaria control, which can improve child health, may have differently affected the areas of high and low malaria endemicity during the time period under study. For these factors to have little influence on the estimated effects, it is sufficient to show that they may not have affected child health differently in the low and high endemicity areas. Such potential confounders include the following factors. First the program geared toward the improvement of the national health system may have affected child health; but it is unlikely that it would have varied in its impact on the low and high malaria endemicity areas (MoHSW 2008, TMIERG 2012). Second, for the last few decades, there has been sustained economic growth, suggesting a subsequent improvement in health indicators. However, there is no indication that this growth would have affected low and high endemicity areas differently (Macro 2005, 2011, MoHSW 2008, TMIERG 2012). Third, the haemophilus influenza vaccination and the use of vitamin A supplementation have been introduced during the time period covered by this paper. Yet, the available reports do not suggest any differences in uptake between the low and high endemicity areas (NMCP-MoHSW 2013).

2.9. Cost-Benefit Analysis

Tanzania's malaria program has harnessed increased attention because of its potential role in improving child health (MoHSW 2008). The justification for the malaria control interventions mainly rests on an equity argument due to the fact that the deleterious effects of malaria are especially acute among the poor who can afford neither the costs of adequate prevention nor those related to the treatment of acute cases of Malaria (Cohn 1972, Hanson et al. 2004). The malaria control interventions can potentially affect economic, demographic, and health outcomes in Africa. Research has shown that a decline in morbidity and an improvement in the quality of life are directly associated with a reduced incidence of malaria (Cohn 1972). Some researchers have hypothesized two possible economic effects of malaria control. First, the improvement in health, stemming from a reduction in malaria prevalence, may increase average hours worked along with productivity during the hours spent at work (Cohn 1972). Second, reduced private and public spending on malaria is thought to free up resources that can be used elsewhere in an economy (Cohn 1972). Because the assumption of full employment is implausible in Tanzania, the first economic effect is likely to be negligible, while the second economic effect is the more plausible to operate. In terms of demographic effects, previous findings have shown that malaria control tends to cause a reduction in mortality due to less exposure to malaria, especially among children and pregnant women, reducing the prevalence of anemia, pre-term birth, low birthweight, and miscarriages (Cohn 1972, Newman 1977). Importantly, a combination of these effects may result in increased fertility. The remainder of this section aims to assess empirically the economic and survival benefits from malaria control in mainland Tanzania.

Despite the controversies surrounding the application of methods related to cost-benefit analysis in the field of health intervention, these methods can provide unique insights into the rules of allocation of social resources (Brent 2007, Layard and Glaister 1994, McIntosh 2010, Mills, Lubell, and Hanson 2008). Because I am primarily interested in child health, I focus on health benefits accruing to children under five as a result of malaria control. In terms of program benefits, I mostly use survival gains and gains in life expectancy at birth estimated from the Cox regression specified in Equation (2.3). The number of lives saved is then computed by multiplying the survival gains by the mean annual number of live births between 2005 and 2010. Population data are drawn from publications from the National Statistical Services of Tanzania. I use the number of live births and the annual growth rate in 2012 to calculate the number of live births and children under five in Tanzania between 2005 and 2010.[5] The mean number of births between 2004 and 2010 and the mean number of children under five between 2005 and 2010 represent the population data used for the valuation of the malaria program's benefits (number of lives saved and years of life saved). Survival probabilities and gains in years of life are calculated with estimates that combine both sex.[6]

With the survival gains, I compute the improvements in expected years of life by using a life-table approach with mortality data from the World Population Prospects:

[5]Tanzanian National Bureau of Statistics, 2014 and UNICEF at:$https://www.unicef.org/infobycountry/tanzania_statistics.html$.

[6]The combined estimates rely on a weighted average of males and females and account for sex ratio at birth.

The 2015 Revision and from the United Nations Population Fund's tools for demographic estimation[7]. Using the modified logit and the Brass logit models (Brass 1971, Murray et al. 2003), I compute the contribution of malaria control to improvement in life expectancy at birth[8]. I choose the United Nations Chile model as the standard because it has emerged as the best fitting model, as measured by the R-squared, from a series of linear regressions of the logit of the survival probabilities from the 2010 UN life table of Tanzania and each of the nine standard life tables available (Coale, Demeny, and Vaughan 2013). I use survival information from 2010 to estimate the gains in life expectancy. To calculate the contribution of malaria control, I subtract the survival gains attributable to malaria control from the under-five survival probabilities prevailing in 2010; then, I assume that there is no improvement in adult survival (people aged between 15 and 60 years old) attributable to malaria control between 2005 and 2010. This latter assumption reflects an extreme case scenario and will estimate a lower bound to the gains in the expected years of life (See Appendix for further details).

To compute the economic benefits of malaria control, I use the value a statistical life from a paper by Leon and Miguel. The value of a statistical life tends to vary by the level of economic development, and estimates are usually higher for the developed countries compared to the developing ones. In this paper, it is then important to use the value of a statistical life computed in Africa. Fortunately, in the paper by Leon and Miguel, the value of a statistical life was derived from a revealed preference approach with data from a survey in Sierra Leone. Specifically, Leon and Miguel obtain the value of a statistical life by assessing the trade-off in mortality risks and travel costs

[7] Retrieved at: http://demographicestimation.iussp.org/content/fitting-model-life-tables-pair-estimates-child-and-adult-mortality.

[8] The estimates produced do not correct for high mortality due to the HIV/AIDS pandemic in Tanzania.

that passengers, who fly through Freetown's Lungi International Airport in 2012, have made. The authors exploit the fact that the capital city is reached by crossing an estuary through either a ferry, helicopter, hovercraft, or water taxi. Because each transportation medium entails different costs, durations, and mortality risks (based on past accidents), Leon and Miguel have been able to investigate how travelers from different countries value their lives. They find that the value of a statistical life among African travelers is about 577,000 US dollars (Leon and Miguel 2013). The value of a statistical life is a key measure for assessing the benefits of health interventions (Brent 2007, Layard and Glaister 1994, McIntosh 2010). It varies widely depending on the valuation methods and the national income level. The VSL used in this analysis is recorded in 2011 PPP dollars (Leon and Miguel 2013). In addition, I calculate a value of a statistical year of life (VOLY), using the VSL and survival probabilities with the methodological approach recommended by the European Union (Lindhjem et al. 2012). The derivation of the VOLY is as follows:

$$VOLY = VSL * \frac{1}{\left(\sum_{t=0}^{T} P_{t,0}(1+r)^{-t}\right)} \tag{2.8}$$

Where:

1. $P_{t,0}$ is the conditional probability of surviving from age 0 to t;
2. r is the social discount rate;
3. T is the maximal age in the population.

As far as the program costs are concerned, I obtain the annual program costs, which provide crude estimates due to the fact that the malaria program is not solely geared to children under five. These cost data do not, unfortunately, include indirect government spending on malaria and other costs related to facility maintenance. I retrieve cost data from official publications on the program budgets (TMIERG 2012) from 2000 to 2010. To assess the costs of malaria program, I assume that the annual program costs pertaining to the year 2010 will prevail until malaria eradication has been achieved because costs have been increasing from 2004 to 2010. To determine the duration of the interventions until complete eradication, I borrow the information about the time span until malaria eradication from the experiences in India and Zanzibar, where eradication is achieved within 10 and 15 years of sustained interventions, respectively (Cohn 1972, MoHSW 2011). Adding a five-year margin to the Zanzibar experience, I assume that mainland Tanzania could achieve malaria eradication within 20 years of malaria control interventions.

Finally, I use a set of three discount rates. The World Bank often relies on discount rates as high as 10 percent in developing countries (World-Bank 2010). In this section, discount rates, ranging from 10 and 20 percent, are used to calculate the net present values of the malaria program. The upper bound is selected in a somewhat arbitrary fashion; but because high discount rates reduce an intervention's chance of passing the net positive present value, such high rates lead to more conservative acceptance decisions. Because cost and benefit data pertain to year 2010 and 2011, I make no inflation adjustment. I compute the net present value of the malaria program by assuming that benefits will continue indefinitely into the future (50 years), while costs will be incurred for 20 years.

Table 2.9 shows the survival outcomes of the malaria control interventions in mainland Tanzania between 2000-5 and 2005-10 based on the regression results reported in Table 2.5 indicating that malaria control interventions have helped avert 17.9 deaths for every 1,000 live births in Tanzania. When it comes to changes in life expectancy, the modified logit method indicates gains of 2.9 and 3.0 years of life in the male and female population, respectively, attributable to malaria control interventions, while the Brass logit method indicates gains of 3.2 and 3.4 in the years of life among males and females respectively. Because these estimated gains in life expectancy from the two methods are fairly similar, in the remainder of this section, I use the results from the modified logit method, which tend to be more precise[9] than those from the Brass logit method. Noting that these estimates provide average program effects, I make use of population data to derive the benefits from the interventions.

Table 2.10 presents population estimates, using the 2012 census data and the intercensal population growth rate. This approach suggests no change in fertility resulting from malaria control interventions. This is a conservative assumption as fertility tends to increase in places where malaria eradication has been achieved (Cohn 1972). The population estimates provide the size of the under-five population and the number of live births in mainland Tanzania, which I use to estimate the social benefits of the program. Between 2005 and 2010, the mean annual number of births in (mainland) Tanzania is 1,682,905. There are, on average, 6,267,694 children under five.

Table 2.11 shows both the annual and discounted benefits of malaria control interventions in mainland Tanzania. The table shows that there are, at the population level, 30,102 deaths averted per year among children under five. This value still falls short of the program target, setting out to reduce the 80,000 malaria related deaths by 80 percent (MoHSW 2008). Panel A of Table 2.11 presents the main ingredients and the output of the cost-benefit analysis, using the value of a statistical life as the

[9] http://demographicestimation.iussp.org/content/fitting-model-life-tables-pair-estimates-child-and-adult-mortality

measure of the program benefit. At discount rates ranging from 10 to 20 percent, the program passes the test of a positive net present value. The net benefits of the malaria control program outweigh the program costs by a factor greater than 120. In sheer numbers, the discounted benefits range from 86.8 trillion (20 discount rate) to 172.2 trillion US dollars (10 percent discount rate), while the discounted costs vary between 0.7 and 1.2 trillion US dollars.

The results mentioned above are robust to changes in the parameter values of the cost-benefit analysis. For example, the malaria program would still have been socially desirable, even if the current program costs had experienced a 100-fold increase. Alternatively, even if the VSL had been as low as 4,600 US dollars the malaria program would have passed the test of a positive net present value at all values of discount rates between 10 and 20 percent. Furthermore, substantial increases in the duration of the intervention, which are equivalent to increases in the number of years for which the program costs are incurred, do not affect the conclusion that the malaria program is highly desirable socially. Finally, discount rates as high as 50 percent still yield a positive net present value. Further analysis with an alternative measure of program benefits strengthens these results.

Panel B of Table 2.11 presents a cost-benefit analysis with the statistical value of a year of life. This analysis focuses on the assessment of the survival implications of malaria control for years of life lived in mainland Tanzania. The results indicate highly positive effects of malaria control, which tend to be larger in magnitude than those from the analysis based on lives saved by the malaria program. Similar to the analysis in Panel A, various sensitivity analyses with changes in the program costs, the value of a statistical year of life, the time horizon of the intervention, and the discount rates do not alter the conclusions about the social desirability of the malaria control interventions.

2.10. Conclusion

Malaria imposes a high burden on child health in Tanzania, where it is thought to account for a large fraction of deaths among children under-five. However, most studies on the impact of malaria on child health at the country level do not provide causal estimates of the impact of the national malaria program on improvement in child health that can be quantified. Exploiting differences in exposure to malaria due to geography in mainland Tanzania, I use a difference-in-difference strategy that provides estimates that plausibly reflect the causal impact of malaria control on child health, after adjusting for the children's socioeconomic characteristics. Consistent with previous randomized control trials in several African countries, I find that mortality has rapidly declined in the high malaria endemicity areas compared with the low malaria endemicity areas because of the malaria control interventions (Alonso et al. 1993, Alonso et al. 1991, Binka et al. 1996, D'Alessandro et al. 1995, Nevill et al. 1996). The estimates suggest a reduction of 17.9 deaths on average for every 1,000 live births attributable to malaria control interventions. In relative terms, the interventions have contributed to 57.7 percent of the reduction in mortality between 2004 and 2010. These estimates are within the bounds of past findings from previous randomized control trials (Alonso et al. 1993, Alonso et al. 1991, D'Alessandro et al. 1995, Fraser-Hurt et al. 1999, Schellenberg et al. 2001).

The study further uncovers other health benefits, stemming from the malaria control interventions. They have caused anemia to drop by 52 percent. The interventions have also helped reduced the prevalence of stunting and low weight-for-age by 36 and 33 percent, respectively.

In light of the large health benefits for children found in this paper, there is justification continuing the large scale interventions in order to strengthen current gains. Assessing the relevance of the malaria control program in a cost-benefit analysis, I find large benefits that outweigh most conservative measures of program costs and discount rates. The malaria program is a highly desirable intervention with high health benefits.

2.11. References

Alexandre, M. A. A., Benzecry, S. G., Siqueira, A. M., Vitor-Silva, S., Melo, G. C., Monteiro, W. M., Alecrim, M. d. G. C. (2015). The association between nutritional status and malaria in children from a rural community in the Amazonian Region: a longitudinal study. PLoS Negl Trop Dis, 9(4), e0003743.

Allison, P. D. (1996). Fixed-effects partial likelihood for repeated events. Sociological Methods & Research, 25(2), 207-222.

Allison, P. D. (2014). Event history and survival analysis: Regression for longitudinal event data (Vol. 46): Sage Publications.

Alonso, P., Lindsay, S., Schellenberg, J. A., Keita, K., Gomez, P., Shenton, F., Cham, K. (1993). A malaria control trial using insecticide-treated bed nets and targeted chemoprophylaxis in a rural area of The Gambia, West Africa: 6. The impact of the interventions on mortality and morbidity from malaria. Transactions of the Royal Society of Tropical Medicine and Hygiene, 87, 37-44.

Alonso, P. L., Lindsay, S. W., Armstrong, J., de Francisco, A., Shenton, F., Greenwood, B., David, P. (1991). The effect of insecticide-treated bed nets on mortality of Gambian children. The Lancet, 337(8756), 1499-1502.

Amemiya, T. (1985). Advanced econometrics: Harvard university press.

Angrist, J. D., & Pischke, J.-S. (2008). Mostly harmless econometrics: An empiricist's companion: Princeton university press.

Aregawi, M., Cibulskis, R. E., Otten, M., & Williams, R. (2009). World malaria report 2009: World Health Organization.

Arinaitwe, E., Gasasira, A., Verret, W., Homsy, J., Wanzira, H., Kakuru, A., Kamya, M. R. (2012). The association between malnutrition and the incidence of malaria among young HIV-infected and-uninfected Ugandan children: a prospective study.

Malaria journal, 11(1), 1.

Bayoh, M., & Lindsay, S. (2003). Effect of temperature on the development of the aquatic stages of Anopheles gambiae sensu stricto (Diptera: Culicidae). Bulletin of entomological research, 93(05), 375-381.

Bayoh, M. N. (2001). Studies on the development and survival of Anopheles gambiae sensu stricto at various temperatures and relative humidities. Durham University.

Bayoh, M. N., & Lindsay, S. W. (2004). Temperaturerelated duration of aquatic stages of the Afrotropical malaria vector mosquito Anopheles gambiae in the laboratory. Medical and veterinary entomology, 18(2), 174-179.

Becher, H., Muller, O., Jahn, A., Gbangou, A., Kynast-Wolf, G., & Kouyat, B. (2004). Risk factors of infant and child mortality in rural Burkina Faso. Bulletin of the World Health Organization, 82(4), 265-273.

Behrman, J. R., & Deolalikar, A. B. (1988). Health and nutrition. Handbook of development economics, 1, 631-711.

Berger, U., Fahrmeir, L., & Klasen, S. (2002). Dynamic modelling of child Mortality in developing countries: Application for Zambia.

Binka, F. N., Kubaje, A., Adjuik, M., Williams, L. A., Lengeler, C., Maude, G., Smith, P. G. (1996). Impact of permethrin impregnated bednets on child mortality in KassenaNankana District, Ghana: a randomized controlled trial. Tropical Medicine & International Health, 1(2), 147-154.

Black, R. E., Cousens, S., Johnson, H. L., Lawn, J. E., Rudan, I., Bassani, D. G., Cibulskis, R. (2010). Global, regional, and national causes of child mortality in 2008: a systematic analysis. The Lancet, 375(9730), 1969-1987.

Bonner, K., Mwita, A., McElroy, P. D., Omari, S., Mzava, A., Lengeler, C., Mtung'e, R. (2011). Design, implementation and evaluation of a national campaign to distribute nine million free LLINs to children under five years of age in Tanzania. Malaria journal, 10(1), 1.

Borucka, J. (2014). Methods for Handling Tied Events in the Cox Proportional Hazard Model. Studia Oeconomica Posnaniensia, 2(2), 91.

Brass, W. (1971). On the scale of mortality. Biological aspects of demography, 69-110.

Brent, R. J. (2007). Applied cost-benefit analysis: Edward Elgar Publishing.

Buffet, P. A., Safeukui, I., Deplaine, G., Brousse, V., Prendki, V., Thellier, M., Mercereau-Puijalon, O. (2011). The pathogenesis of Plasmodium falciparum malaria in humans: insights from splenic physiology. Blood, 117(2), 381-392.

Centers for Disease Control. (2010). Anopheles mosquitoes. See http://www. cdc. gov/malaria/about/biology/mosquitoes/index. html.

Clark, I. A., & Cowden, W. B. (2003). The pathophysiology of falciparum malaria. Pharmacology & therapeutics, 99(2), 221-260.

Cleves, M. (2010). An introduction to survival analysis using Stata: Stata Press.

Coale, A. J., Demeny, P., & Vaughan, B. (2013). Regional Model Life Tables and Stable Populations: Studies in Population: Elsevier.

Cohn, E. J. (1972). Assessment of malaria eradication costs and benefits. The American journal of tropical medicine and hygiene, 21(5 Suppl), 663-667.

D'Alessandro, U., Olaleye, B., Langerock, P., Aikins, M., Thomson, M., Cham, M., Cham, B. (1995). Mortality and morbidity from malaria in Gambian children after introduction of an impregnated bednet programme. The Lancet, 345(8948), 479-483.

Danquah, I., Dietz, E., Zanger, P., Reither, K., Ziniel, P., Bienzle, U., & Mockenhaupt, F. P. (2009). Reduced efficacy of intermittent preventive treatment of malaria in malnourished children. Antimicrobial agents and chemotherapy, 53(5), 1753-1759.

De Onis, M., Blssner, M., & Borghi, E. (2012). Prevalence and trends of stunting among pre-school children, 19902020. Public health nutrition, 15(01), 142-148.

Deen, J., Walraven, G., & Von Seidlein, L. (2002). Increased risk for malaria in chronically malnourished children under 5 years of age in rural Gambia. Journal of tropical pediatrics, 48(2), 78-83.

Efron, B. (1977). The efficiency of Cox's likelihood function for censored data. Journal of the American Statistical Association, 72(359), 557-565.

Ehrhardt, S., Burchard, G. D., Mantel, C., Cramer, J. P., Kaiser, S., Kubo, M., Mockenhaupt, F. P. (2006). Malaria, Anemia, and Malnutrition in African Children: Defining Intervention Priorities. The Journal of Infectious Diseases, 194(1), 108-114.

Elo, I. T., & Miller, J. E. (1991). Ever-breastfed status and health at birth.

Eswarappa, S. M., Estrela, S., & Brown, S. P. (2012). Within-host dynamics of multi-species infections: facilitation, competition and virulence. PloS one, 7(6), e38730.

Ferreira, E. d. A., Alexandre, M. A., Salinas, J. L., de Siqueira, A. M., Benzecry, S. G., de Lacerda, M. V., & Monteiro, W. M. (2015). Association between anthropometry-based nutritional status and malaria: a systematic review of observational studies. Malaria journal, 14(1), 1-23.

Fillol, F., Cournil, A., Boulanger, D., Ciss, B., Sokhna, C., Targett, G., Simondon, K. B. (2009). Influence of wasting and stunting at the onset of the rainy season on subsequent malaria morbidity among rural preschool children in Senegal. The American journal of tropical medicine and hygiene, 80(2), 202-208.

Fraser-Hurt, N., Felger, I., Edoh, D., Steiger, S., Mashaka, M., Masanja, H., Beck, H.-P. (1999). 9. Effect of insecticide-treated bed nets on haemoglobin values, prevalence and multiplicity of infection with Plasmodium falciparum in a randomized controlled trial in Tanzania. Transactions of the Royal Society of Tropical Medicine and Hygiene, 93(Supplement 1), 47-51.

Friedman, J. F., Kurtis, J. D., Ramadhan, M., Opollo, M., Lanar, D. E., & Duffy, P. E. (2003). Malaria is related to decreased nutritional status among male adolescents and adults in the setting of intense perennial transmission. Journal of Infectious Diseases, 188(3), 449-457.

Friedman, M. (1982). Piecewise exponential models for survival data with covariates. The Annals of Statistics, 101-113.

Gamble, C., Ekwaru, P. J., Garner, P., & Ter Kuile, F. O. (2007). Insecticide-treated nets for the prevention of malaria in pregnancy: a systematic review of randomised controlled trials. PLoS Med, 4(3), e107.

Genton, B., Al-Yaman, F., Alpers, M., & Mokela, D. (1997). Indicators of fatal outcome in paediatric cerebral malaria: a study of 134 comatose Papua New Guinean children. International journal of epidemiology, 26(3), 670-676.

Genton, B., Al-Yaman, F., Ginny, M., Taraika, J., & Alpers, M. P. (1998). Relation of anthropometry to malaria morbidity and immunity in Papua New Guinean children. The American journal of clinical nutrition, 68(3), 734-741.

Greenwood, B. M., Bradley, A., Greenwood, A., Byass, P., Jammeh, K., Marsh, K., Hayes, R. (1987). Mortality and morbidity from malaria among children in a rural area of The Gambia, West Africa. Transactions of the Royal Society of Tropical Medicine and Hygiene, 81(3), 478-486.

Guerrant, R. L., DeBoer, M. D., Moore, S. R., Scharf, R. J., & Lima, A. A. (2013). The impoverished guta triple burden of diarrhoea, stunting and chronic disease. Nature Reviews Gastroenterology and Hepatology, 10(4), 220-229.

Guerrant, R. L., Ori, R. B., Moore, S. R., Ori, M. O., & Lima, A. A. (2008). Malnutrition as an enteric infectious disease with long-term effects on child development. Nutrition reviews, 66(9), 487-505.

Guo, G., & Rodriguez, G. (1992). Estimating a multivariate proportional hazards model for clustered data using the EM algorithm, with an application to child survival in Guatemala. Journal of the American Statistical Association, 87(420), 969-976.

Guyatt, H. L., & Snow, R. W. (2001). Malaria in pregnancy as an indirect cause of infant mortality in sub-Saharan Africa. Transactions of the Royal Society of Tropical Medicine and Hygiene, 95(6), 569-576.

Hall, N., Karras, M., Raine, J. D., Carlton, J. M., Kooij, T. W., Berriman, M., Christophides, G. K. (2005). A comprehensive survey of the Plasmodium life cycle by genomic, transcriptomic, and proteomic analyses. Science, 307(5706), 82-86.

Hanson, K., Goodman, C., Lines, J., Meek, S., Bradley, D., & Mills, A. (2004). The economics of malaria control interventions.

Hanson, K., Nathan, R., Marchant, T., Mponda, H., Jones, C., Bruce, J., Schellenberg, J. A. (2008). Vouchers for scaling up insecticide-treated nets in Tanzania: methods for monitoring and evaluation of a national health system intervention. BMC Public Health, 8(1), 1.

Hautvast, J. L., Tolboom, J. J., Kafwembe, E. M., Musonda, R. M., Mwanakasale, V., van Staveren, W. A., Monnens, L. A. (2000). Severe linear growth retardation in rural Zambian children: the influence of biological variables. The American journal of clinical nutrition, 71(2), 550-559.

Hoddinott, J., Maluccio, J. A., Behrman, J. R., Flores, R., & Martorell, R. (2008). Effect of a nutrition intervention during early childhood on economic productivity in Guatemalan adults. The Lancet, 371(9610), 411-416.

Hsieh, F. Y. (1995). A cautionary note on the analysis of extreme data with Cox regression. The American Statistician, 49(2), 226-228.

Kabula, B., Derua, Y. A., Tungu, P. K., Massue, D. J., Sambu, E., Mosha, F. W., & Kisinza, W. N. (2011). Malaria entomological profile in Tanzania from 1950 to 2010: a review of mosquito distribution, vectorial capacity and insecticide resistance. Tanzania journal of health research, 13(5).

Kang, H., Kreuels, B., Adjei, O., Krumkamp, R., May, J., & Small, D. S. (2013).

The causal effect of malaria on stunting: a Mendelian randomization and matching approach. International journal of epidemiology, 42(5), 1390-1398.

Kirby, M., & Lindsay, S. (2004). Responses of adult mosquitoes of two sibling species, Anopheles arabiensis and A. gambiae ss (Diptera: Culicidae), to high temperatures. Bulletin of entomological research, 94(05), 441-448.

Klein, J. P., Van Houwelingen, H. C., Ibrahim, J. G., & Scheike, T. H. (2016). Handbook of survival analysis: CRC Press.

Korenromp, E. L., Williams, B. G., Gouws, E., Dye, C., & Snow, R. W. (2003). Measurement of trends in childhood malaria mortality in Africa: an assessment of progress toward targets based on verbal autopsy. The Lancet infectious diseases, 3(6), 349-358.

Layard, R., & Glaister, S. (1994). Cost-benefit analysis: Cambridge University Press.

Len, G., & Miguel, E. (2013). Transportation choices and the value of statistical life. Retrieved from Lindhjem, H., Navrud, S., Biausque, V., & Braathen, N. (2012). Mortality risk valuation in environment, health and transport policies: OECD, Paris.

Macro, I. I. (2005). Tanzania demographic and health survey 2004-2005.

Macro, I. I. (2011). Tanzania demographic and health survey 2010.

Malaria, R. B. (2005). World malaria report 2005. World Health Organization and UNICEF.

Malaria Roll Back World Health Organization. (2000). The Abuja Declaration and the Plan of Action. An Extract from the African Summit on Roll Back Malaria, Abuja, 25.

Man, W., Weber, M., Palmer, A., Schneider, G., Wadda, R., Jaffar, S., Greenwood, B. (1998). Nutritional status of children admitted to hospital with different diseases and its relationship to outcome in The Gambia, West Africa. Tropical Medicine & International Health, 3(8), 678-686.

Masanja, H., de Savigny, D., Smithson, P., Schellenberg, J., John, T., Mbuya, C., Smith, T. (2008). Child survival gains in Tanzania: analysis of data from demographic and health surveys. The Lancet, 371(9620), 1276-1283.

McIntosh, E. (2010). Applied methods of cost-benefit analysis in health care (Vol. 4): Oxford University Press.

McSweeney, C., New, M., & Lizcano, G. (2002). UNDP Climate Change Country Profiles Tanzania.

Mills, A., Lubell, Y., & Hanson, K. (2008). Malaria eradication: the economic, financial and institutional challenge. Malaria journal, 7(1), 1.

Mitangala, P., DAlessandro, U., Donnen, P., Hennart, P., Porignon, D., Balaluka, G. B., Wilmet, M. D. (2013). Infection palustre et tat nutritionnel: rsultats dune cohorte denfants gs de 6 59 mois au Kivu en Rpublique dmocratique du Congo. Revue d'pidmiologie et de sant publique, 61(2), 111-120.

Mockenhaupt, F. P., Ehrhardt, S., Burkhardt, J., Bosomtwe, S. Y., Laryea, S., Anemana, S. D., Gellert, S. (2004). Manifestation and outcome of severe malaria in children in northern Ghana. The American journal of tropical medicine and hygiene, 71(2), 167-172.

Mohammed, A., Mandike, R., Molteni, F., Chaki, P., Okumu, F., & Njau, R. (2015). An epidemiological profile of malaria and its control in mainland Tanzania. National Malaria Control Programme, WHO-Tanzania, Ifakara Health Institute, KEMRI-Wellcome Trust.

MoHSW. (2003). National Malaria Medium Term Strategic Plan 2002 - 2007.

MoHSW. (2008). Medium Term Malaria Strategic Plan 2008 - 2013.

MoHSW. (2011). Zanzibar Malaria Program Performance Review.

Mosley, W. H., & Chen, L. C. (1984). An analytical framework for the study of child survival in developing countries. Population and development review, 10, 25-45.

Msukwak, A. A. (1994). Child Anthropometry and Mortality In Malawi: Testing for Effect Modification by Age and Length of Follow-up and Confounding by Socioeconomic Factors.

Muller, O., Garenne, M., Kouyat, B., & Becher, H. (2003). The association between proteinenergy malnutrition, malaria morbidity and allcause mortality in West African children. Tropical Medicine & International Health, 8(6), 507-511.

Murray, C. J., Ferguson, B. D., Lopez, A. D., Guillot, M., Salomon, J. A., & Ahmad, O. (2003). Modified logit life table system: principles, empirical validation, and application. Population Studies, 57(2), 165-182.

Nevill, C., Some, E., Mung'Ala, V., Muterni, W., New, L., Marsh, K., Snow, R. (1996). Insecticidetreated bednets reduce mortality and severe morbidity from malaria among children on the Kenyan coast. Tropical Medicine & International Health, 1(2), 139-146.

Newman, P. (1977). Malaria and mortality. Journal of the American Statistical Association, 72(358), 257-263.

Newman, R. D. (2012). World Malaria Report 2011.

Newton, C. R., & Krishna, S. (1998). Severe falciparum malaria in children: current understanding of pathophysiology and supportive treatment. Pharmacology & therapeutics, 79(1), 1-53.

NMCP-MoHSW. (2013). An epidemiological profile of malaria and its control in Mainland Tanzania.

Nyakeriga, A. M., Troye-Blomberg, M., Chemtai, A. K., Marsh, K., & Williams, T. N. (2004). Malaria and nutritional status in children living on the coast of Kenya. The American journal of clinical nutrition, 80(6), 1604-1610.

Olney, D. K., Kariger, P. K., Stoltzfus, R. J., Khalfan, S. S., Ali, N. S., Tielsch, J. M., Pollitt, E. (2013). Developmental effects of micronutrient supplementation and malaria in zanzibari children. Early human development, 89(9), 667-674.

Olumese, P., Sodeinde, O., Ademowo, O., & Walker, O. (1997). Protein energy malnutrition and cerebral malaria in Nigerian children. Journal of tropical pediatrics, 43(4), 217-219.

O'Meara, W. P., Mangeni, J. N., Steketee, R., & Greenwood, B. (2010). Changes in the burden of malaria in sub-Saharan Africa. The Lancet infectious diseases, 10(8), 545-555.

Organization, W. H. (2015). Malaria rapid diagnostic test performance: results of WHO product testing of malaria RDTs: round 6 (2014-2015).

Padonou, G., Le Port, A., Cottrell, G., Guerra, J., Choudat, I., Rachas, A., Martin-Prevel, Y. (2014). Factors associated with growth patterns from birth to 18 months in a Beninese cohort of children. Acta tropica, 135, 1-9.

Pazzaglia, G., & Woodward, W. E. (1982). An analysis of the relationship of host factors to clinical falciparum malaria by multiple regression techniques. The American journal of tropical medicine and hygiene, 31(2), 202-210.

Rabe-Hesketh, S., & Skrondal, A. (2008). Multilevel and longitudinal modeling using Stata: STATA press.

Renaudin, P. (1997). Evaluation de l'tat nutritionnel de l'enfant de moins de 5 ans Moundou, Tchad: Relations avec la morbidit et la mortalit hospitalires. Mdecine tropicale, 57(1), 49-54.

Renggli, S., Mandike, R., Kramer, K., Patrick, F., Brown, N. J., McElroy, P. D., Nathan, R. (2013). Design, implementation and evaluation of a national campaign to deliver 18 million free long-lasting insecticidal nets to uncovered sleeping spaces in Tanzania. Malaria journal, 12(1), 1.

Rosenzweig, M. R., & Schultz, T. P. (1983). Estimating a household production function: Heterogeneity, the demand for health inputs, and their effects on birth weight.

The Journal of Political Economy, 723-746.

Rosenzweig, M. R., & Wolpin, K. I. (1988). Heterogeneity, intrafamily distribution, and child health. Journal of human resources, 437-461.

Rowe, A. K., Steketee, R. W., Arnold, F., Wardlaw, T., Basu, S., Bakyaita, N., Cibulskis, R. E. (2007). Viewpoint: Evaluating the impact of malaria control efforts on mortality in subSaharan Africa. Tropical Medicine & International Health, 12(12), 1524-1539.

Schellenberg, D., Menendez, C., Kahigwa, E., Font, F., Galindo, C., Acosta, C., Urassa, H. (1999). African children with malaria in an area of intense Plasmodium falciparum transmission: features on admission to the hospital and risk factors for death. The American journal of tropical medicine and hygiene, 61(3), 431-438.

Schellenberg, J. R. A., Abdulla, S., Nathan, R., Mukasa, O., Marchant, T. J., Kikumbih, N., Mshinda, H. (2001). Effect of large-scale social marketing of insecticide-treated nets on child survival in rural Tanzania. The Lancet, 357(9264), 1241-1247.

Schultz, T. P. (1999). Health and schooling investments in Africa. The Journal of Economic Perspectives, 13(3), 67-88.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference: Houghton, Mifflin and Company.

Shankar, A. H. (2000). Nutritional modulation of malaria morbidity and mortality. Journal of Infectious Diseases, 182(Supplement 1), S37-S53.

Shanks, G. D., Hay, S. I., & Bradley, D. J. (2008). Malaria's indirect contribution to all-cause mortality in the Andaman Islands during the colonial era. The Lancet infectious diseases, 8(9), 564-570.

Sinka, M. E., Bangs, M. J., Manguin, S., Coetzee, M., Mbogo, C. M., Hemingway, J., Kabaria, C. W. (2010). The dominant Anopheles vectors of human malaria in Africa, Europe and the Middle East: occurrence data, distribution maps and bionomic prcis. Parasites & vectors, 3(1), 1.

Smithson, P., Florey, L., Salgado, S. R., Hershey, C. L., Masanja, H., Bhattarai, A., Group, T. M. I. E. R. (2015). Impact of malaria control on mortality and anemia among tanzanian children less than five years of age, 19992010. PloS one, 10(11), e0141112.

Snow, R., Byass, P., Shenton, F., & Greenwood, B. (1991). The relationship between anthropometric measurements and measurements of iron status and susceptibility to malaria in Gambian children. Transactions of the Royal Society of Tropical Medicine and Hygiene, 85(5), 584-589.

Snow, R. W., Trape, J.-F., & Marsh, K. (2001). The past, present and future of childhood malaria mortality in Africa. TRENDS in Parasitology, 17(12), 593-597.

Sowunmi, A., Gbotosho, G. O., Adedeji, A. A., Fateye, B. A., Sabitu, M. F., Happi, C. T., & Fehintola, F. A. (2007). Effects of acute Plasmodium falciparum malaria on body weight in children in an endemic area. Parasitology research, 101(2), 343-349.

Steketee, R. W., & Campbell, C. C. (2010). Impact of national malaria control scale-up programmes in Africa: magnitude and attribution of effects. Malaria journal, 9(1), 1.

Strauss, J., & Thomas, D. (1998). Health, nutrition, and economic development. Journal of economic literature, 36(2), 766-817.

Taylor, P., & Mutambu, S. (1986). A review of the malaria situation in Zimbabwe with special reference to the period 19721981. Transactions of the Royal Society of Tropical Medicine and Hygiene, 80(1), 12-19.

Ter Kuile, F. O., Terlouw, D. J., Phillips-Howard, P. A., Hawley, W. A., Friedman, J. F., Kolczak, M. S., Vulule, J. M. (2003). Impact of permethrin-treated bed nets on malaria and all-cause morbidity in young children in an area of intense perennial malaria transmission in western Kenya: cross-sectional survey. The American journal of tropical medicine and hygiene, 68(4 suppl), 100-107.

Therneau, T. M., & Grambsch, P. M. (2000). Modeling survival data: extending the Cox model: Springer Science & Business Media.

TMIERG. (2012). Evaluation of the Impact of Malaria Interventions on Mortality in Children in Mainland Tanzania.

Todd, P. E. (2007). Evaluating social programs with endogenous program placement and selection of the treated. Handbook of development economics, 4, 3847-3894.

Tonglet, R. (1999). How useful are anthropometric, clinical and dietary measurements of nutritional status as predictors of morbidity of young children in central Africa? Tropical Medicine & International Health, 4(2), 120-130.

Trussell, J., & Hammerslough, C. (1983). A hazards-model analysis of the covariates of infant and child mortality in Sri Lanka. Demography, 20(1), 1-26.

Van den Berg, G. J. (2001). Duration models: specification, identification and multiple durations. Handbook of econometrics, 5, 3381-3460.

Van den Broeck, J., Eeckels, R., & Vuylsteke, J. (1993). Influence of nutritional status on child mortality in rural Zaire. The Lancet, 341(8859), 1491-1495.

Vaupel, J. W., Manton, K. G., & Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography, 16(3), 439-454.

WHO. (2008). Medium-term strategic plan 2008-2013 Amended-draft.

WHO. (2014). World malaria report 2013: World Health Organization.

Williams, T., Maitland, K., Phelps, L., Bennett, S., Peto, T., Viji, J., Bowden, D. (1997). Plasmodium vivax: a cause of malnutrition in young children. Qjm, 90(12), 751-757.

World Health Organization. (2000). Severe falciparum malaria. Transactions of the Royal Society of Tropical Medicine and Hygiene, 94, Supplement 1, 1-90.

World-Bank. (2010). Cost - Benefit Analysis in World Bank Projects.

2.12. Tables

Table 2.1: Distribution of Children under Five Ever Born to Women by Year of Birth and Survival Status

Panel A: 2004 Demographic and Health Survey				
	Alive	Deceased	Total	
Year of birth	N	N	N	Col %
1999	44	1	45	0.7
2000	1102	167	1269	19.3
2001	1166	124	1290	19.6
2002	1200	110	1310	19.9
2003	1251	96	1347	20.5
2004	1240	57	1297	19.7
2005	26	1	27	0.4
Total	6029	556	6585	100.0

Panel B: 2010 Demographic and Health Survey				
	Alive	Deceased	Total	
Year of birth	N	N	N	Col %
2005	825	66	891	14.8
2006	1167	80	1247	20.7
2007	1097	93	1190	19.7
2008	1153	72	1225	20.3
2009	1159	61	1220	20.2
2010	247	7	254	4.2
Total	5648	379	6027	100.0

Note: Sample characteristics are based on unweighted data
Source: DHS, waves 2004-5 and 2009-10

Table 2.2: Test of Proportional-Hazards Assumption Based on the Shoenfeld residual

Control variables	Correlation coefficient (rho)	chi2	df	Prob>chi2
High malaria endemicity (prevalence>= 5%)*Survey year=2009-10	0.035	1.120	1	0.290
High malaria endemicity (prevalence>= 5%)	0.014	0.190	1	0.666
Indicator of survey year=2009-10	-0.048	2.180	1	0.140
Child's sex (male=1)	-0.051	2.450	1	0.118
Season of birth (rainy s.=1)	0.047	2.030	1	0.155
Whether mother attended antenatal care during pregnancy	0.119	13.980	1	0.000
Mother's age (categorical)				
<20	0.004	0.010	1	0.904
30-39	0.014	0.180	1	0.672
40-49	0.013	0.160	1	0.693
Mother's education				
primary	0.054	2.740	1	0.098
secondary or higher	0.007	0.040	1	0.834
Education of mother's partner				
primary	0.019	0.360	1	0.548
secondary or higher	-0.007	0.050	1	0.820
Sex of household head (male=1)	-0.101	9.490	1	0.002
Whether child lives in urban areas	0.007	0.050	1	0.829
Whether mother reports distance as a big concern for her use of healthcare	-0.065	3.990	1	0.046
Whether household has clean drinking water	-0.057	3.050	1	0.081
Whether household has improved sanitation	-0.044	1.830	1	0.177
Whether walls are made from modern materials	-0.029	0.760	1	0.382
Whether floor is made from modern materials	-0.052	2.530	1	0.112
Whether roof is made from modern materials	0.009	0.070	1	0.788
Global test		56.590	21	0.000

Source: DHS, waves 2004-5 and 2009-10

Table 2.3: Relationship between measures of child health and population-adjusted falciparum prevalence (PFPR) in children aged 2 to 10 years

Panel A: Unadjusted hazard ratios of a Cox regression of child survival on PFPR

	(1)	(2)	(3)	(4)
	2004		2010	
PFPR (continuous)	1.00		1.01*	
	[0.00]		[0.00]	
Indicator of on PFPR				
<5% [ref.]				
5 - 10 %		1.68*		0.53*
		[0.36]		[0.13]
10 - 50 %		2.13**		1.18
		[0.42]		[0.21]
>50 %		1.88**		0.89
		[0.46]		[0.23]
N	6,585	6,585	6,027	6,027

Panel B: Unadjusted coefficientsof a linear regression of hemoglobin density (anemia) on PFPR

	2004		2010	
PFPR (continuous)	-0.16**		-0.12**	
	[0.01]		[0.01]	
Indicator of on PFPR				
<5% [ref.]				
5 - 10 %		0.82		-5.91**
		[1.08]		[0.91]
10 - 50 %		-4.20**		-9.43**
		[0.92]		[0.78]
>50 %		-7.37**		-10.31**
		[1.24]		[1.10]
N				

Panel C: Unadjusted coefficients of a linear regression of z-scores for height-for-age on PFPR

	2004		2010	
PFPR (continuous)	0.01**		0.00**	
	[0.00]		[0.00]	
Indicator of on PFPR				
<5% [ref.]				
5 - 10 %		-0.16		-0.38**
		[0.10]		[0.08]
10 - 50 %		0.10		-0.23**
		[0.08]		[0.07]
>50 %		0.48**		0.01
		[0.11]		[0.10]
N	4,675	4,675	5,031	5,031

Panel D: Unadjusted coefficients of a linear regression of z-scores for weight-for-age on PFPR

	2004		2010	
PFPR (continuous)	0.00**		0.00**	
	[0.00]		[0.00]	
Indicator of on PFPR				
<5% [ref.]				
5 - 10 %		-0.18**		-0.28**
		[0.07]		[0.06]
10 - 50 %		0.00		-0.08
		[0.06]		[0.06]
>50 %		0.16*		0.07
		[0.08]		[0.08]
N	4,675	4,675	5,031	5,031

Standard errors in brackets
** p<0.01, * p<0.05, + p<0.1
Source: DHS, waves 2004-5 and 2009-10

Table 2.4: Sample Characteristics by Survey Year and Malaria Endemicity (Mean or Proportion), Children <5

	Pooled sample	2004		2010	
		Low endemicity	High endemicity	Low endemicity	High endemicity
Panel A: Program coverage					
Proportion of households with treated bednets	0.37	0.15	0.13	0.50	0.63
Proportion of households with bednets, where all children sleep under bednets	0.44	0.22	0.23	0.56	0.67
N	$11,677^{a}$	596	5,433	521	5,127
Panel B: Morbidity and Nutritional Health					
Proportion with moderate anemia (7.0-9.9g/dL)	0.36	0.26	0.45	0.24	0.28
	$[N = 10,463]^{b}$	[540]	[4,870]	[473]	[4,580]
Weight (z-score)	0.00	-0.03	-0.02	0.08	0.02
	[0.01]	[0.04]	[0.02]	[0.05]	[0.01]
Height (z-score)	0.00	0.00	-0.04	0.06	0.04
	[0.01]	[0.04]	[0.01]	[0.05]	[0.02]
Proportion stunted	0.48	0.42	0.50	0.49	0.46
Proportion with low weight-for-age	0.18	0.15	0.18	0.20	0.17
N	$9,706^{c}$	507	4,424	443	4,232
Panel C: Socio-Demographic Characteristics					
Proportion male	0.50	0.50	0.50	0.52	0.49
Child's age at the time of the survey or age at death (months)	29.25	28.61	29.25	30.19	29.24
	[0.15]	[0.67]	[0.22]	[0.72]	[0.23]
Proportion born in rainy season	0.60	0.55	0.61	0.56	0.61
Whether mother attended antenatal care during pregnancy	0.60	0.66	0.61	0.68	0.58
Mother's age (years)	29.09	28.51	28.77	30.00	29.42
	[0.06]	[0.25]	[0.09]	[0.27]	[0.09]
Mother's age (categorical)					
<20	0.04	0.04	0.05	0.02	0.04
20-29	0.54	0.57	0.56	0.50	0.52
30-39	0.39	0.42	0.38	0.46	0.40
40-49	0.11	0.07	0.10	0.13	0.12
Mother's education					
no education	0.27	0.13	0.29	0.13	0.27
primary	0.70	0.80	0.68	0.81	0.69
secondary or higher	0.04	0.07	0.03	0.06	0.04
Education of mother's partner					
no education	0.18	0.10	0.20	0.11	0.18
primary	0.75	0.80	0.74	0.76	0.75
secondary or higher	0.07	0.10	0.06	0.13	0.06
Proportion of male headed households	0.84	0.85	0.84	0.88	0.84
Whether child lives in uraban areas	0.16	0.21	0.16	0.27	0.15
Whether mother reports distance as a big concern for her use of healthcare	0.35	0.46	0.43	0.25	0.26
Whether household has clean drinking water	0.56	0.57	0.46	0.60	0.67
Whether household has improved sanitation	0.78	0.84	0.78	0.85	0.77
Whether walls are made from modern materials	0.59	0.54	0.56	0.56	0.64
Whether floor is made from modern materials	0.17	0.27	0.14	0.33	0.18
Whether roof is made from modern materials	0.43	0.61	0.35	0.72	0.47
N	12,612	624	5,961	555	5,472

Standard errors in brackets
([a]) Sample excludes deceased children and children under 6 months
([b]) Sample excludes deceased children, children under 6 months, and cases with missing data on weight and height

Source: DHS, waves 2004-5 and 2009-10

Table 2.5: Hazard Ratios from Cox Regression of Child Mortality, Children <5

		(1) Under-5 mortality	(2) Under-5 mortality	(3) Under-5 mortality
Main Effects				
High malaria endemicity (prevalence>=5%)*Survey year=2009-10		0.51*	0.49**	0.49**
		[0.14]	[0.13]	[0.13]
High malaria endemicity (prevalence>=5%)		2.02**	1.90**	1.91**
		[0.39]	[0.37]	[0.37]
Indicator of survey year=2009-10		1.37	1.44	1.44
		[0.35]	[0.37]	[0.37]
Child's sex (male=1)			1.23**	1.23**
			[0.08]	[0.08]
Season of birth (rainy s.=1)			0.86*	0.86*
			[0.06]	[0.06]
Whether mother attended antenatal care during pregnancy			0.41**	0.40**
			[0.03]	[0.03]
Mother's age (categorical)				
	20-29 (ref.)			
	<20		1.23	1.25
			[0.22]	[0.22]
	30-39		0.91	0.90
			[0.06]	[0.06]
	40-49		1.09	1.09
			[0.12]	[0.12]
Mother's education				
	no education (ref.)			
	primary		0.98	0.98
			[0.08]	[0.08]
	secondary or higher		0.76	0.76
			[0.18]	[0.18]
Education of mother's partner				
	no education (ref.)			
	primary		0.95	0.95
			[0.09]	[0.09]
	secondary or higher		1.15	1.15
			[0.19]	[0.19]
Sex of household head (male=1)			0.82*	0.88
			[0.07]	[0.08]
Whether child lives in urban areas			1.20+	1.20+
			[0.13]	[0.13]
Whether mother reports distance as a big concern for her use of healthcare			1.00	1.00
			[0.07]	[0.07]
Whether household has clean drinking water			0.95	0.95
			[0.06]	[0.06]
Whether household has improved sanitation			1.21*	1.21*
			[0.11]	[0.10]
Whether walls are made from modern materials			0.96	0.96
			[0.07]	[0.07]
Whether floor is made from modern materials			1.12	1.12
			[0.13]	[0.13]
Whether roof is made from modern materials			0.84*	0.84*
			[0.07]	[0.07]
				.
Time-varying interactions				
Whether mother attended antenatal care during pregnancy				1.08**
				[0.03]
Sex of household head (male=1)				0.89**
				[0.03]
Whether mother reports distance as a big concern for her use of healthcare				0.98
				[0.03]
N		12,612	12,612	12,612

Standard errors in brackets
** p<0.01, * p<0.05, + p<0.1

Source: DHS, waves 2004-5 and 2009-10

Table 2.6: Odds Ratios from Logistic Regression of Indicators of Anemia and Nutritional Health, Children <5

	(1) Anemia (7.0-9.9g/dL)	(2) Stunting (<2SD)	(3) Low weight-for-age (<2SD)	(4) Anemia (7.0-9.9g/dL)	(5) Stunting (<2SD)	(6) Low weight-for-age (<2SD)
High malaria endemicity (prevalence>=5%)* Survey year=2009-10	0.53**	0.64**	0.65*	0.48**	0.64**	0.67*
	[0.08]	[0.09]	[0.12]	[0.07]	[0.09]	[0.12]
High malaria endemicity (prevalence>= 5%)	2.33**	1.35**	1.28+	2.29**	1.15	1.12
	[0.24]	[0.13]	[0.17]	[0.24]	[0.11]	[0.15]
Indicator of survey year=2009-10	0.91	1.34*	1.45*	0.99	1.39*	1.52*
	[0.13]	[0.18]	[0.25]	[0.15]	[0.18]	[0.27]
Child's age (months)				0.96**	1.01**	1.00
				[0.00]	[0.00]	[0.00]
Child's sex (male=1)				1.21**	1.28**	1.29**
				[0.05]	[0.05]	[0.07]
Season of birth (rainy s.=1)				1.04	1.11*	1.06
				[0.05]	[0.05]	[0.06]
Whether mother attended antenatal care during pregnancy				1.02	0.97	0.97
				[0.05]	[0.05]	[0.06]
Mother's age (categorical)						
20-29 (ref.)						
<20				1.02	1.26*	1.15
				[0.12]	[0.15]	[0.17]
30-39				1.03	0.98	1.07
				[0.05]	[0.04]	[0.06]
40-49				1.01	1.03	0.98
				[0.07]	[0.07]	[0.09]
Mother's education						
no education (ref.)						
primary				0.90+	0.96	0.82**
				[0.05]	[0.05]	[0.05]
secondary or higher				0.98	0.62**	0.60*
				[0.13]	[0.09]	[0.13]
Education of mother's partner						
no education (ref.)						
primary				0.85**	1.04	0.93
				[0.05]	[0.06]	[0.07]
secondary or higher				0.75*	0.71**	0.58**
				[0.09]	[0.08]	[0.10]
Sex of household head (male=1)				1.16*	0.80**	0.87+
				[0.07]	[0.05]	[0.06]
Whether child lives in urban areas				0.98	0.83**	0.97
				[0.07]	[0.06]	[0.09]
Whether mother reports distance as a big concern for her use of healthcare				0.98	0.90*	0.98
				[0.05]	[0.04]	[0.06]
Whether household has clean drinking water				1.04	0.92+	0.85**
				[0.05]	[0.04]	[0.05]
Whether household has improved sanitation				0.92	1.21**	1.22**
				[0.05]	[0.07]	[0.09]
Whether walls are made from modern materials				1.02	1.04	0.78**
				[0.05]	[0.05]	[0.05]
Whether floor is made from modern materials				0.83*	0.67**	0.74**
				[0.06]	[0.05]	[0.08]
Whether roof is made from modern materials				0.84**	0.84**	0.86*
				[0.05]	[0.04]	[0.06]
N	10,463[a]	9,706[b]	9,706[b]	10,463[a]	9,706[b]	9,706[b]

Robust SE in brackets
** p<0.01, * p<0.05, + p<0.1
([a]) Sample excludes deceased children and children under 6 months
([b]) Sample excludes deceased children,children under 6 months, and cases with missing data on weight and height

Source: DHS, waves 2004-5 and 2009-10

Table 2.7: Hazard Ratios of Cox Regression and Odds Ratios from Logistic Regression, Extending Previous Models, Children <5

	(1) Odds Ratios	(2)	(3) Hazard Ratios	(4)
	Under-5 mort.	Under-5 mort. (No breast-feed.)	Under-5 mort. (breast-feed.)	Under-5 mort. (FE)
High malaria endemicity (prevalence>= 5%)* Survey year=2009-10	0.48**	0.56*	0.55*	0.52*
	[0.13]	[0.15]	[0.15]	[0.14]
High malaria endemicity (prevalence>=5%)	1.92**	1.84**	1.84**	1.86**
	[0.38]	[0.36]	[0.36]	[0.38]
Indicator of survey year=2009-10	1.46	1.24	1.25	1.35
	[0.37]	[0.33]	[0.34]	[0.35]
Child's sex (male=1)	1.23**	1.21**	1.21**	1.23**
	[0.08]	[0.08]	[0.08]	[0.08]
Season of birth (rainy s.=1)	0.86*	0.88+	0.88*	0.85*
	[0.06]	[0.06]	[0.06]	[0.06]
Whether mother attended antenatal care during pregnancy	0.41**	0.43**	0.44**	0.40**
	[0.03]	[0.03]	[0.03]	[0.03]
Mother's age (categorical)				
20-29 (ref.)				
<20	1.25	1.23	1.21	1.29
	[0.22]	[0.22]	[0.22]	[0.23]
30-39	0.91	0.89	0.90	0.91
	[0.06]	[0.06]	[0.06]	[0.06]
40-49	1.10	1.10	1.12	1.13
	[0.12]	[0.12]	[0.12]	[0.12]
Mother's education				
no education (ref.)				
primary	0.98	1.02	1.01	0.96
	[0.08]	[0.08]	[0.08]	[0.08]
secondary or higher	0.76	0.81	0.81	0.73
	[0.18]	[0.19]	[0.19]	[0.18]
Education of mother's partner				
no education (ref.)				
primary	0.95	0.93	0.93	0.99
	[0.09]	[0.08]	[0.08]	[0.09]
secondary or higher	1.15	1.08	1.09	1.18
	[0.19]	[0.18]	[0.18]	[0.20]
Sex of household head (male=1)	0.82*	0.84*	0.84*	0.87
	[0.07]	[0.07]	[0.07]	[0.08]
Whether child lives in urban areas	1.21+	1.25*	1.24*	0.98
	[0.13]	[0.13]	[0.13]	[0.16]
Whether mother reports distance as a big concern for her use of healthcare	1.00	1.00	1.00	0.97
	[0.07]	[0.07]	[0.07]	[0.07]
Whether household has clean drinking water	0.95	0.94	0.93	0.94
	[0.06]	[0.06]	[0.06]	[0.07]
Whether household has improved sanitation	1.21*	1.18*	1.20*	1.20*
	[0.11]	[0.10]	[0.11]	[0.11]
Whether walls are made from modern materials	0.96	0.96	0.96	1.01
	[0.07]	[0.07]	[0.07]	[0.08]
Whether floor is made from modern materials	1.12	1.08	1.08	1.11
	[0.13]	[0.13]	[0.13]	[0.14]
Whether roof is made from modern materials	0.84*	0.83*	0.83*	0.82*
	[0.07]	[0.07]	[0.07]	[0.07]
Child's age	0.89**			
	[0.00]			
Breastfeeding[a]			0.54**	
			[0.07]	
Interaction with time[b]	No	No	Yes	Yes
N	352, 225[c]	352, 225[d]	352, 225[d]	12,612

Robust SE in brackets
** p<0.01, * p<0.05, + p<0.1
([a]) Takes on the value one during the months when child is breastfed and zero otherwise
([b]) LR test of theta=0: chibar2(01) = 20.3
([c]) Sample consists of multiple records for each child
([d]) Sample consists of multiple records for each child and excludes cases with missing data on breastfeeding

Source: DHS, waves 2004-5 and 2009-10

Table 2.8: Hazard Ratios of Cox Regression and Odds Ratios from Placebo Regression of Whether Child Has Anemia and Hazard Ratios from Placebo Cox Regression for Child Mortality, Children <5

	(1) Under-5 mort. (Hazard Ratios)	(2) Stunting (<2SD)	(3) Low weight-for--age (<2SD)
High malaria endemicity (prevalence>=5%)* Survey year=2004-05	1.51	1.10	1.16
	[0.44]	[0.22]	[0.28]
High malaria endemicity (prevalence>=5%)	1.16	1.11	1.02
	[0.27]	[0.19]	[0.21]
Indicator of survey year=2004-05	0.72	0.75	0.59*
	[0.21]	[0.14]	[0.14]
Child's age (months)		1.01**	1.00
		[0.00]	[0.00]
Child's sex (male=1)	1.1	1.25**	1.18**
	[0.07]	[0.06]	[0.07]
Season of birth (rainy s.=1)	0.96	1.00	1.08
	[0.07]	[0.05]	[0.07]
Whether mother attended antenatal care during pregnancy	0.32**	1.07	0.95
	[0.04]	[0.07]	[0.08]
Mother's age (categorical)			
20-29 (ref.)			
<20	1.50**	1.31*	1.58**
	[0.22]	[0.15]	[0.21]
30-39	0.9	0.96	0.98
	[0.07]	[0.05]	[0.06]
40-49	1.05	1.06	1.09
	[0.12]	[0.09]	[0.11]
Mother's education			
no education (ref.)			
primary	0.95	0.98	0.84*
	[0.08]	[0.06]	[0.06]
secondary or higher	0.73	0.70*	0.60*
	[0.17]	[0.11]	[0.14]
Sex of household head (male=1)	0.82+	0.73**	0.86+
	[0.09]	[0.05]	[0.07]
Whether child lives in urban areas	1	0.80**	0.84+
	[0.11]	[0.06]	[0.08]
Whether household has clean drinking water	0.99	0.95	1.00
	[0.07]	[0.05]	[0.06]
Whether household has improved sanitation	1.11	1.23**	1.21*
	[0.10]	[0.08]	[0.09]
Whether floor is made from modern materials	0.97	0.47**	0.58**
	[0.12]	[0.04]	[0.06]
Interaction with time	Yes	No	No
N	9,329	6,957[a]	6,957[a]

Robust SE in brackets
** p<0.01, * p<0.05, + p<0.1
([a]) Sample excludes deceased children, children under 6 months, and cases with missing data on weight and height

Source: DHS, waves 1999 and 2004-5

Table 2.9: Survival Outcomes of Malaria Control between 2000-5 and 2005-10 in Mainland Tanzania

Outcome		**2005**	**2010 - With malaria program**	**2010 - Without malaria program**	**Contribution of malaria**
Under five survivorship (x1000)		112.0 [a]	81.0 [a]	98.9 [b]	17.9 [c]
Life Expectancy at birth					
Modified logit [c]					
	Male	61.1	65.6	62.7	2.9
	Female	65.7	69.8	66.8	3.0
	Combined	63.3	67.6	64.7	2.9
Brass Logit [c]					
	Male	61.7	68.1	64.9	3.2
	Female	68.0	75.3	71.9	3.4
	Combined	64.8	71.6	68.3	3.3

([a]) 2011 DHS Country Report (Tanzania)
([b]) Estimates based on Table 2.5 and 2011 DHS Country Report (Tanzania)
([c]) World Population Prospects: The 2015 Revision - Life Table of the United Republic of Tanzania.
The United Nation Chile is the Standard

Table 2.10: Population Estimates in Mainland Tanzania

Population Characteristics	Year	Number of Births Both Sexes	Population Under-5 Male	Female
Population Data	2012	1,898,300	3,535,673	3,534,222
Annual Intercensal Growth rates (2002-2012)		0.027	0.027	0.027
Period Mean		1,682,905	3,134,490	3,133,204
Estimated Population				
	2010	1,798,510	3,349,810	3,348,435
	2009	1,750,600	3,260,575	3,259,237
	2008	1,703,966	3,173,718	3,172,415
	2007	1,658,575	3,089,174	3,087,906
	2006	1,614,393	3,006,882	3,005,648
	2005	1,571,387	2,926,782	2,925,581

Source: Tanzanian National Bureau of Statistics, 2014 and UNICEF at:
https://www.unicef.org/infobycountry/tanzania_statistics.html

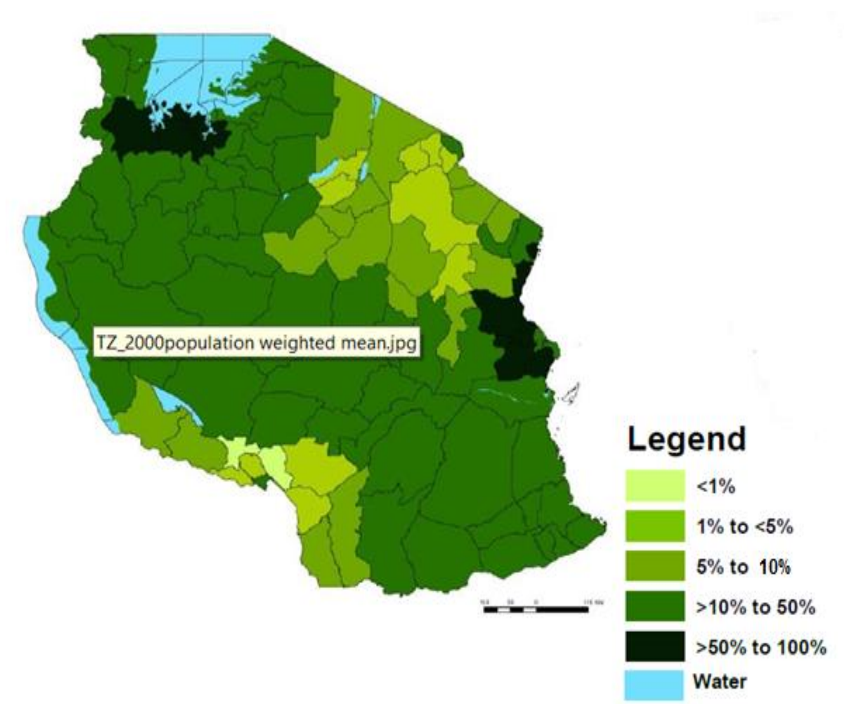
Table 2.11: Economic and Demographic Net Benefits from the Malaria Program (Cost of Program for 20 years)

Panel A: Economic Net Benefits Based on Number of Deaths Averted						
Social Discount Rate (percent)	**Value of Statistical Life ($US)**	**Number of Under-5 Deaths Averted**	**Annual Program**		**Discounted**[b]	
			Costs	**Benefit**	**Costs**	**Benefit**
			($US millions)[a]			
10	577,000	30,102.80	140	17,369.32	1191.90	172,213.54
15	577,000	30,102.80	140	17,369.32	876.31	115,688.58
20	577,000	30,102.80	140	17,369.32	681.74	86,837.03
Panel B: Economic Benefits Based on Gains in Years of Life						
Social Discount Rate (percent)	**Value of Year of Life ($US)**[c]	**Total Years of Life Gained**[d]	**Annual Program**		**Discounted**	
			Costs	**Benefit**	**Costs**	**Benefit**
			($US millions)			
10	247,735	17,228,515.15	45	4,268,098.68	383.11	42,317,406.62
15	334,943	17,228,515.15	45	5,770,575.07	281.67	38,434,999.83
20	411,065	17,228,515.15	45	7,082,037.08	219.13	35,406,294.37

([a]) Costs do not include government indirect spending on malaria, e.g. staff time and infrastructure maintenance
([b]) Costs are discounted over 20 years and benefits over 50 years
([c]) Based on Equation (2.8)
([d]) Based on contribution of malaria to life expectancy at birth (Table 2.9) times population under five (Table 2.10)
Source: Author's Own Calculations

2.13. Figures

Figure 2.1: Population-Adjusted Mean Prevalence of Malaria among Children Aged 2 -10 in Tanzania



Source: An Epidemiological Profile of Malaria and Its Control in Mainland Tanzania (NMCP-MoHSW 2013), p. 75

Figure 2.2: Kaplan-Meier Estimates of Survivor among Children under - 5,Tanzania (2004 and 2010)

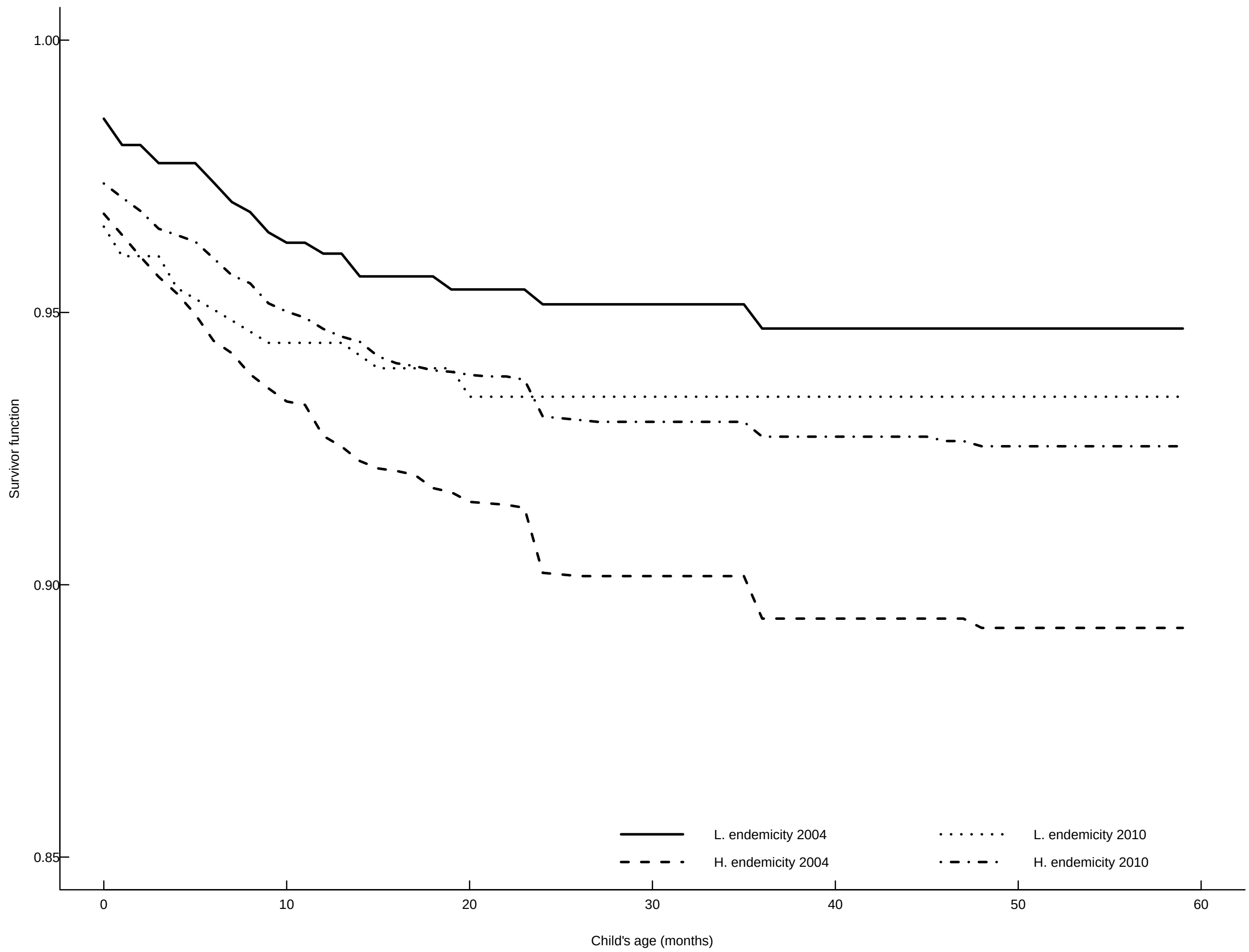
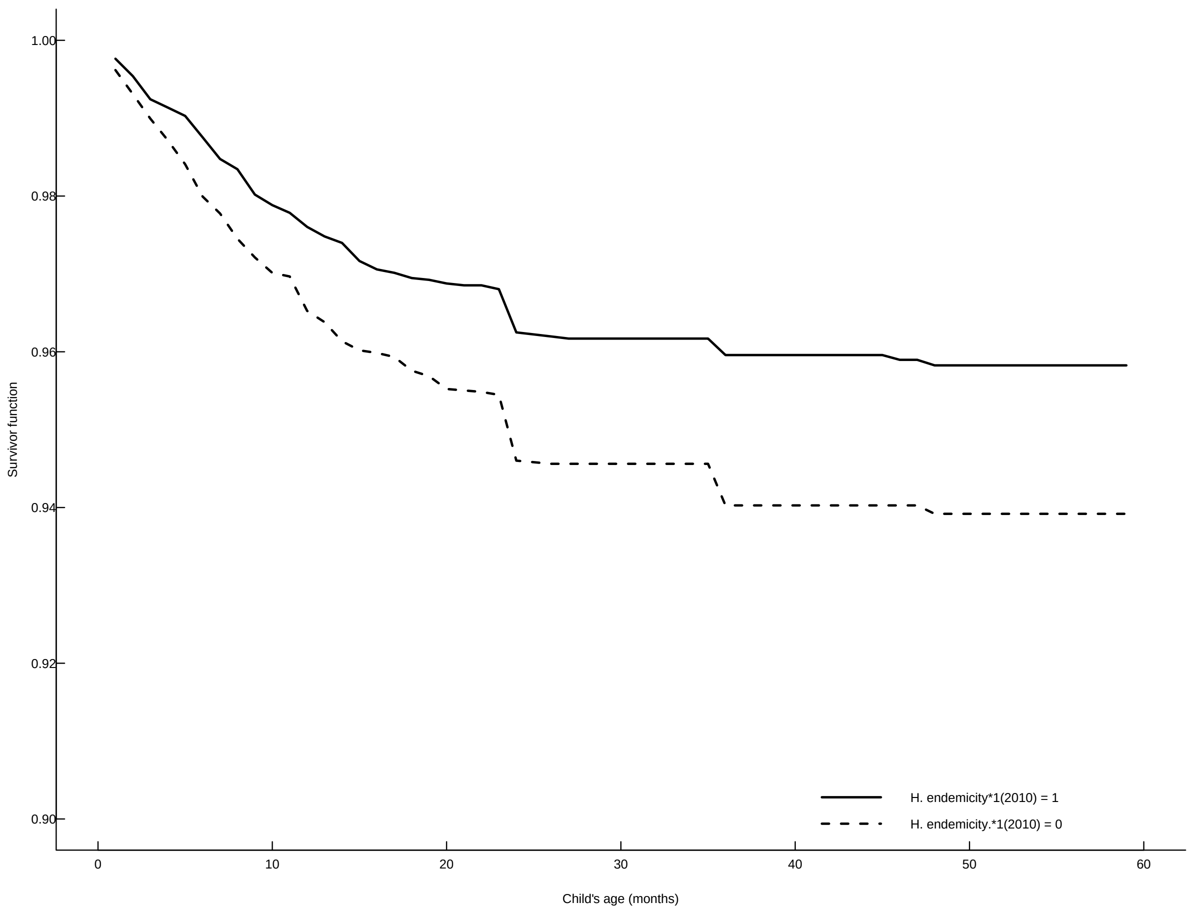


Figure 2.3: Estimated Survivor: Effects of Malaria Program,Tanzania (2004 and 2010)



2.14. Appendix

I estimated a single value of life expectancy at birth in 2005, while, in 2010, two values of life expectancy at birth are computed to reflect the two possible scenarios implied by the absence or presence of the malaria interventions. In other words, in 2010, a first scenario reflects changes in child mortality in the presence of the malaria program, and the second refers to the case where the program would have been absent. To produce life expectancy at birth, I relied on a two-step approach. First, I chose a model life table among Coale-Demeny and UN life tables (Coale, Demeny, and Vaughan 2013).[10] The preferred model life table was the one with the highest statistical correlation with the 2005 Tanzanian life table available in the World Population Prospects: The 2015 Revision.[11] In subsequent step, I estimated the value of life expectancy at birth using the Brass logit and the Modified Logit approaches, using the program made available.[12] For these relational models, the estimation of life expectancy is achieved with two sets of survival probabilities (one at younger ages and the other at adult ages). Consistent with the demography literature, I used survival probabilities ${}_5q_0$ and ${}_{15}q_{45}$.

Table A2.1 reports the R-squared of the regression of the 2005 male and female life tables on their counterparts among the model life tables. It shows that the UN Chile model is the preferred model life table based on my selection criterion (largest correlation for males and second largest for females).

Table A2.2 presents the survival probabilities for ${}_5q_0$ and ${}_{15}q_{45}$. In the counterfactual scenario estimated for 2010, I further assumed that the survival probability

[10] Princeton North, Princeton South, Princeton East, Princeton West, United Nations General, UN Latin America, UN Chile, UN South Asia, and UN Far East.

[11] https://esa.un.org/unpd/wpp/Download/Standard/Mortality/

[12] http://demographicestimation.iussp.org/content/fitting-model-life-tables-pair-estimates-child-and-adult-mortality.

between 15 and 60 years would have remained unchanged at the level of 2005. This assumption would most likely result in underestimation of the program benefit in terms of gains in life expectancy.

With this model life table and the survival probabilities, I produced the estimates of life expectancy from three life tables for each sex and for the years 2005 and 2010 (two life tables). The results are shown below in Tables A2.3 through A2.8. The results for each sex are combined in a weighted average to implement the cost benefit analysis.

Table A2.1: R-squared of Regression of the 2005 Tanzanian Life Table on Model Life Tables

	Adjusted R-squared	
Model Life Table	Male	Female
UN Chile	0.915	0.881
UN Far East	0.913	0.887
UN General	0.904	0.871
UN Latin America	0.897	0.864
Princeton West	0.896	0.868
Princeton East	0.890	0.853
UN South Asia	0.887	0.849
Princeton North	0.886	0.857
Princeton South	0.867	0.831

Source: World Population Prospects: The 2015 Revision

Table A2.2: Survival Probabilities for Calculating Life Expectancy

	2005[a]		2010 - With Malaria Program[a]		2010 - Without Malaria Program [b]	
	Male	Female	Male	Female	Male	Female
$_{0}q_{5}$	0.11209704	0.09900963	0.08132011	0.07291554	0.10045011	0.09204554
$_{15}q_{45}$	0.45012868	0.45240581	0.35379091	0.3484443	0.45012868	0.45240581

([a]) World Population Prospects: The 2015 Revision - Life Table of the United Republic of Tanzania.
([b]) 2010 Life Table of the United Republic of Tanzania, adjusted for survival gains from Table 2.9.

Table A2.3: Estimated Life Table with the Modified Logit Approach - 2005

		Male			Female		
		$_{5}q_{0}$	0.1121		$_{5}q_{0}$	0.0990	
		$_{45}q_{15}$	0.4501		$_{45}q_{15}$	0.4524	
		Y(5)	-1.0347		Y(5)	-1.1041	
		a	-0.0002		a	-0.0305	
		β	0.6505	-0.2239	β	0.6386	-0.3051
		e0	61.1		e0	65.7	
		Fitted	Fitted life table		Fitted	Fitted life table	
Age (x)	n	logits	l(x)	nmx	logits	l(x)	nmx
0	1		1.0000	0.09024		1.0000	0.07779
1	4	-1.1694	0.9120	0.00673	-1.2477	0.9238	0.00628
5	5	-1.0347	0.8879	0.00179	-1.1041	0.9010	0.00143
10	5	-0.9962	0.8800	0.00096	-1.0691	0.8946	0.00072
15	5	-0.9766	0.8758	0.00089	-1.0524	0.8914	0.00080
20	5	-0.9588	0.8719	0.00167	-1.0343	0.8878	0.00134
25	5	-0.9271	0.8646	0.00211	-1.0051	0.8819	0.00167
30	5	-0.8894	0.8555	0.00257	-0.9708	0.8745	0.00190
35	5	-0.8464	0.8446	0.00341	-0.9342	0.8663	0.00226
40	5	-0.7939	0.8303	0.00481	-0.8934	0.8565	0.00286
45	5	-0.7268	0.8106	0.00721	-0.8455	0.8444	0.00414
50	5	-0.6381	0.7818	0.01104	-0.7825	0.8271	0.00646
55	5	-0.5222	0.7397	0.01747	-0.6955	0.8007	0.01041
60	5	-0.3714	0.6776	0.02497	-0.5764	0.7600	0.01595
65	5	-0.1982	0.5978	0.03890	-0.4276	0.7016	0.02714
70	5	0.0171	0.4915	0.06061	-0.2284	0.6122	0.04625
75	5	0.2845	0.3615	0.09424	0.0304	0.4848	0.07772
80	5	0.6224	0.2236	0.14681	0.3608	0.3271	0.12607
85		1.0724	0.1048	0.22872	0.7865	0.1718	0.20452

Source: http://demographicestimation.iussp.org/content/fitting-model-life-tables-pair--estimates-child-and-adult-mortality

Table A2.4: Estimated Life Table with the Modified Logit Approach - 2010 (Scenario: With Malaria Program)

		Male			Female		
		$_5q_0$	0.0813		$_5q_0$	0.0729	
		$_{45}q_{15}$	0.3538		$_{45}q_{15}$	0.3484	
		Y(5)	-1.2123		Y(5)	-1.2714	
		a	-0.1777		a	-0.1978	
		β	0.6505	-0.1779	β	0.6386	-0.2343
		e0	65.6		e0	69.8	
		Fitted	Fitted life table		Fitted	Fitted life table	
Age (x)	n	logits	l(x)	nmx	logits	l(x)	nmx
0	1		1.0000	0.06610		1.0000	0.05894
1	4	-1.3320	0.9349	0.00438	-1.3921	0.9418	0.00395
5	5	-1.2123	0.9187	0.00120	-1.2714	0.9271	0.00091
10	5	-1.1766	0.9132	0.00073	-1.2411	0.9229	0.00046
15	5	-1.1559	0.9099	0.00092	-1.2262	0.9207	0.00049
20	5	-1.1310	0.9057	0.00132	-1.2109	0.9185	0.00081
25	5	-1.0970	0.8997	0.00145	-1.1865	0.9147	0.00107
30	5	-1.0621	0.8932	0.00167	-1.1561	0.9099	0.00127
35	5	-1.0243	0.8858	0.00221	-1.1218	0.9041	0.00160
40	5	-0.9779	0.8761	0.00327	-1.0817	0.8969	0.00216
45	5	-0.9153	0.8618	0.00526	-1.0316	0.8873	0.00328
50	5	-0.8270	0.8394	0.00854	-0.9632	0.8729	0.00518
55	5	-0.7067	0.8043	0.01397	-0.8693	0.8505	0.00835
60	5	-0.5489	0.7499	0.02103	-0.7436	0.8157	0.01318
65	5	-0.3650	0.6748	0.03340	-0.5862	0.7636	0.02268
70	5	-0.1419	0.5705	0.05304	-0.3803	0.6815	0.03959
75	5	0.1282	0.4363	0.08444	-0.1171	0.5583	0.06868
80	5	0.4619	0.2842	0.13493	0.2139	0.3947	0.11522
85		0.8984	0.1422	0.21561	0.6337	0.2197	0.19329

Source: http://demographicestimation.iussp.org/content/fitting-model-life-tables-pair-
-estimates-child-and-adult-mortality

Table A2.5: Estimated Life Table with the Modified Logit Approach - 2010 (Scenario: Without Malaria Program)

		Male			Female		
		$_5q_0$	0.1005		$_5q_0$	0.0920	
		$_{45}q_{15}$	0.4501		$_{45}q_{15}$	0.4524	
		Y(5)	-1.0961		Y(5)	-1.1445	
		a	-0.0616		a	-0.0709	
		β	0.6505	-0.2421	β	0.6386	-0.3136
		e0	62.7		e0	66.8	
		Fitted	Fitted life table		Fitted	Fitted life table	
Age (x)	n	logits	l(x)	nmx	logits	l(x)	nmx
0	1		1.0000	0.08105		1.0000	0.07276
1	4	-1.2256	0.9206	0.00582	-1.2825	0.9286	0.00563
5	5	-1.0961	0.8995	0.00156	-1.1445	0.9080	0.00129
10	5	-1.0586	0.8926	0.00087	-1.1106	0.9021	0.00065
15	5	-1.0386	0.8887	0.00092	-1.0943	0.8992	0.00071
20	5	-1.0184	0.8846	0.00155	-1.0769	0.8960	0.00119
25	5	-0.9858	0.8778	0.00185	-1.0489	0.8907	0.00150
30	5	-0.9491	0.8697	0.00222	-1.0155	0.8840	0.00173
35	5	-0.9079	0.8601	0.00294	-0.9794	0.8764	0.00208
40	5	-0.8575	0.8475	0.00422	-0.9388	0.8673	0.00268
45	5	-0.7920	0.8298	0.00648	-0.8904	0.8558	0.00392
50	5	-0.7034	0.8033	0.01013	-0.8260	0.8392	0.00614
55	5	-0.5860	0.7635	0.01622	-0.7374	0.8138	0.00989
60	5	-0.4328	0.7038	0.02361	-0.6167	0.7744	0.01527
65	5	-0.2558	0.6252	0.03702	-0.4658	0.7174	0.02604
70	5	-0.0379	0.5189	0.05804	-0.2650	0.6295	0.04464
75	5	0.2305	0.3868	0.09095	-0.0051	0.5026	0.07558
80	5	0.5669	0.2435	0.14286	0.3253	0.3428	0.12356
85		1.0123	0.1167	0.22441	0.7497	0.1825	0.20201

Source: http://demographicestimation.iussp.org/content/fitting-model-life-tables-pair-
-estimates-child-and-adult-mortality

Table A2.6: Estimated Life Table with the Brass Logit Approach - 2005

		Male			Female		
		$_5q_0$	0.1121		$_5q_0$	0.0990	
		$_{45}q_{15}$	0.4501		$_{45}q_{15}$	0.4524	
		Y(5)	-1.0347		Y(5)	-1.1041	
		a	-0.1625		a	-0.4190	
		β	0.7954	-0.2286	β	0.6941	-0.3260
		e0	63.1		e0	70.7	
		Standard	Fitted life table		Standard	Fitted life table	
Age (x)	n	logits	l(x)	nmx	logits	l(x)	nmx
0	1		1.0000	0.10155		1.0000	0.08691
1	4	-1.1882	0.9016	0.00384	-1.1106	0.9153	0.00394
5	5	-1.0966	0.8879	0.00093	-0.9871	0.9010	0.00075
10	5	-1.0710	0.8838	0.00078	-0.9601	0.8976	0.00059
15	5	-1.0503	0.8804	0.00137	-0.9396	0.8949	0.00096
20	5	-1.0151	0.8743	0.00209	-0.9072	0.8906	0.00137
25	5	-0.9647	0.8653	0.00261	-0.8633	0.8846	0.00165
30	5	-0.9062	0.8540	0.00325	-0.8134	0.8773	0.00198
35	5	-0.8392	0.8402	0.00411	-0.7572	0.8687	0.00240
40	5	-0.7623	0.8231	0.00544	-0.6940	0.8583	0.00295
45	5	-0.6712	0.8010	0.00727	-0.6222	0.8458	0.00376
50	5	-0.5638	0.7724	0.00990	-0.5385	0.8300	0.00494
55	5	-0.4370	0.7350	0.01397	-0.4395	0.8097	0.00700
60	5	-0.2849	0.6853	0.02005	-0.3158	0.7818	0.01009
65	5	-0.1028	0.6198	0.02902	-0.1623	0.7433	0.01564
70	5	0.1144	0.5357	0.04198	0.0364	0.6873	0.02317
75	5	0.3724	0.4335	0.06089	0.2758	0.6119	0.03547
80	5	0.6811	0.3190	0.09003	0.5686	0.5122	0.05615
85	5	1.0654	0.2026	0.13793	0.9359	0.3867	0.09165
90	5	1.5729	0.1018	0.20913	1.4118	0.2457	0.14909
95	5	2.2637	0.0364	0.31888	2.0449	0.1191	0.24036
100	5	3.2248	0.0081	0.40263	2.9066	0.0393	0.32774
105	5	4.4209	0.0012	0.46642	4.0038	0.0088	0.40140
110		5.7861	0.0001	0.54031	5.3051	0.0015	0.49163

Source: http://demographicestimation.iussp.org/content/fitting-model-life-tables-pair-
-estimates-child-and-adult-mortality

Table A2.7: Estimated Life Table with the Brass Logit Approach - 2010 (Scenario: With Malaria Program)

		Male			Female		
		$_5q_0$	0.0813		$_5q_0$	0.0729	
		$_{45}q_{15}$	0.3538		$_{45}q_{15}$	0.3484	
		Y(5)	-1.2123		Y(5)	-1.2714	
		a	-0.3400		a	-0.5862	
		β	0.7954	-0.1823	β	0.6941	-0.2520
		e0	68.1		e0	75.3	
		Standard	Fitted life table		Standard	Fitted life table	
Age (x)	n	logits	l(x)	nmx	logits	l(x)	nmx
0	1		1.0000	0.07231		1.0000	0.06305
1	4	-1.1882	0.9289	0.00278	-1.1106	0.9379	0.00289
5	5	-1.0966	0.9187	0.00067	-0.9871	0.9271	0.00056
10	5	-1.0710	0.9156	0.00057	-0.9601	0.9245	0.00044
15	5	-1.0503	0.9130	0.00100	-0.9396	0.9225	0.00071
20	5	-1.0151	0.9085	0.00152	-0.9072	0.9192	0.00101
25	5	-0.9647	0.9016	0.00191	-0.8633	0.9146	0.00122
30	5	-0.9062	0.8930	0.00239	-0.8134	0.9090	0.00147
35	5	-0.8392	0.8824	0.00304	-0.7572	0.9023	0.00178
40	5	-0.7623	0.8691	0.00404	-0.6940	0.8943	0.00220
45	5	-0.6712	0.8517	0.00544	-0.6222	0.8845	0.00282
50	5	-0.5638	0.8288	0.00750	-0.5385	0.8721	0.00373
55	5	-0.4370	0.7982	0.01073	-0.4395	0.8560	0.00532
60	5	-0.2849	0.7564	0.01570	-0.3158	0.8335	0.00775
65	5	-0.1028	0.6992	0.02333	-0.1623	0.8018	0.01219
70	5	0.1144	0.6220	0.03489	0.0364	0.7544	0.01844
75	5	0.3724	0.5219	0.05264	0.2758	0.6878	0.02904
80	5	0.6811	0.4005	0.08136	0.5686	0.5946	0.04775
85	5	1.0654	0.2660	0.13010	0.9359	0.4684	0.08164
90	5	1.5729	0.1392	0.20398	1.4118	0.3127	0.13954
95	5	2.2637	0.0511	0.31663	2.0449	0.1589	0.23382
100	5	3.2248	0.0115	0.40214	2.9066	0.0541	0.32532
105	5	4.4209	0.0017	0.46635	4.0038	0.0123	0.40084
110		5.7861	0.0002	0.54081	5.3051	0.0020	0.49390

Source: http://demographicestimation.iussp.org/content/fitting-model-life-tables-pair--estimates-child-and-adult-mortality

Table A2.8: Estimated Life Table with the Brass Logit Approach - 2010 (Scenario: Without Malaria Program)

		Male			Female		
		$_5q_0$	0.1005		$_5q_0$	0.0920	
		$_{45}q_{15}$	0.4501		$_{45}q_{15}$	0.4524	
		Y(5)	-1.0961		Y(5)	-1.1445	
		a	-0.2239		a	-0.4593	
		β	0.7954	-0.2468	β	0.6941	-0.3338
		e0	64.9		e0	71.9	
		Standard	Fitted life table		Standard	Fitted life table	
Age (x)	n	logits	l(x)	nmx	logits	l(x)	nmx
0	1		1.0000	0.09031		1.0000	0.08045
1	4	-1.1882	0.9120	0.00344	-1.1106	0.9213	0.00366
5	5	-1.0966	0.8995	0.00083	-0.9871	0.9080	0.00070
10	5	-1.0710	0.8958	0.00070	-0.9601	0.9048	0.00055
15	5	-1.0503	0.8927	0.00123	-0.9396	0.9023	0.00090
20	5	-1.0151	0.8872	0.00187	-0.9072	0.8982	0.00127
25	5	-0.9647	0.8789	0.00235	-0.8633	0.8925	0.00154
30	5	-0.9062	0.8687	0.00293	-0.8134	0.8857	0.00184
35	5	-0.8392	0.8560	0.00371	-0.7572	0.8776	0.00223
40	5	-0.7623	0.8403	0.00492	-0.6940	0.8678	0.00275
45	5	-0.6712	0.8199	0.00659	-0.6222	0.8560	0.00352
50	5	-0.5638	0.7933	0.00902	-0.5385	0.8411	0.00462
55	5	-0.4370	0.7582	0.01278	-0.4395	0.8218	0.00656
60	5	-0.2849	0.7111	0.01848	-0.3158	0.7953	0.00949
65	5	-0.1028	0.6482	0.02700	-0.1623	0.7584	0.01476
70	5	0.1144	0.5660	0.03952	0.0364	0.7044	0.02198
75	5	0.3724	0.4639	0.05810	0.2758	0.6308	0.03389
80	5	0.6811	0.3462	0.08718	0.5686	0.5323	0.05414
85	5	1.0654	0.2232	0.13543	0.9359	0.4060	0.08934
90	5	1.5729	0.1136	0.20752	1.4118	0.2609	0.14696
95	5	2.2637	0.0410	0.31819	2.0449	0.1279	0.23895
100	5	3.2248	0.0092	0.40248	2.9066	0.0425	0.32722
105	5	4.4209	0.0014	0.46640	4.0038	0.0096	0.40128
110		5.7861	0.0002	0.54047	5.3051	0.0016	0.49211

Source: http://demographicestimation.iussp.org/content/fitting-model-life-tables-pair-
-estimates-child-and-adult-mortality

CHAPTER 3 : Immigration and Household Extension in the United States: New Patterns and Explanations

Abstract

Previous research on family household extension among the foreign-born has mainly focused on the socioeconomic, demographic, and cultural correlates of living arrangements among a subsample of Mexican and Asian immigrants. We extend this research by including all major sending regions; by distinguishing between horizontal and vertical extension; and by accounting for the uneven geographic distribution of immigrants across the country. Drawing on data from the five percent sample of the 2001-2013 waves of the American Community Survey, we show not only large differentials in the prevalence of extension across immigrant groups, but also substantial variation in the type and predictors of extension, and the extent to which these differences with native whites are explained by socio-demographic composition. Overall, traditional theories of extension do a better job of explaining horizontal than vertical extension, and among relatively disadvantaged immigrant groups (i.e., Mexicans and West Indians) than more positively selected groups (i.e., South East Asians and Canadians/Europeans). African immigrants often fall in between these two extremes. We also show that accounting for immigrant concentration in more expensive housing markets explains an important share of the immigrant-native gap in extension, suggesting that previous analyses exaggerated the role of culture in explaining variation in living arrangements.

3.1. Introduction

It is well established that living arrangements vary substantially between native-born and foreign-born populations within the United States, even net of socio-economic and demographic differences across groups (Glick, Bean, and Van Hook 1997, Goldscheider and Bures 2003, Van Hook and Glick 2007). Immigrants' greater propensity to live in extended households is potentially a source of concern, as it may signal economic distress or a lack of incorporation into larger mainstream society (Kamo 2000, Moen and Wethington 1992). A clear understanding of the nature and source of disparate living arrangements across groups is also a critical issue in its own right, as they influence all evaluations and comparisons conducted at the household level; a failure to account for immigrants tendency to live in extended households could result in an underestimation of ethno-racial inequality in a wide variety of economic outcomes, from household income to homeownership.

In spite of the importance of living arrangements to stratification, the literature on the topic suffers from a number of limitations. First, the majority of research on this topic focuses on Latin American (particularly Mexican) and Asian immigrants. The tremendous diversification in the origins of the U.S. immigrant population over the past several decades necessitates a reprisal and re-evaluation of extension patterns with more up-to-date data and a more comprehensive set of immigrant groups. In addition, remarkable heterogeneity across immigrant groups with respect to human capital and background characteristics (Borjas 1994, Camarota 2012, Elo et al. 2015, Feliciano 2005, Jasso 2011, Jasso et al. 2004) offers new leverage in the attempt to separate economic from cultural sources of disparities in household extension of the foreign born from the native born.

Another limitation of the household extension literature is that it often fails to distinguish between different types of extension. As Van Hook and Glick (2007) have highlighted, the determinants of vertical extension (across generations) often differ from those of horizontal extension (within generations). However, at present there is a dearth of research on specific types of extension. As such, we know relatively little about how the particular pattern of within-and across-generation extension may differ across national origin groups, and how the predictors of living arrangements vary by both region of origin and type of extension. Finally, a common conclusion in the literature on immigrant extension is that much of the residual difference in living arrangements that remains between the foreign born and the native born after accounting for socio-demographic characteristics is attributable to disparate cultural preferences. However, previous studies have failed to consider how the uneven geographic distribution of groups across the country may also contribute to these disparities. Specifically, immigrant populations tend to be far more highly concentrated than the native born geographically, and are particularly over-represented in urban and coastal areas, where the cost of living is higher. It is thus important to take into consideration not only individual and household level characteristics, but also differences across groups in local context when assessing variation in extension patterns.

Drawing on data from the 2001 through 2013 waves of the American Community Survey, we aim to address these gaps in our understanding of immigrant household extension. First, we estimate up-to-date prevalence rates of extension among native-born non-Hispanic whites and multiple foreign-born groups (Mexicans, immigrants from the rest of America, West Indians, South East Asians, immigrants from the rest of Asia, Africans, Oceanians, and Europeans), including the relative contribution of socio-demographic characteristics to household extension. Second, we assess the prevalence and predictors of extension separately for horizontal and vertical extension. Third, we investigate the extent to which variation in local context helps explain some of the residual differences across groups, net of socio-demographic differences, evident in previous studies. And finally, we examine variation in the predictors of extension across immigrant groups, separately for vertical and horizontal extension. Results offer new insights into the social forces structuring household extension, and the differences in family structure between immigrants and the native born.

3.2. Theoretical Background

3.2.1. Economic theories of extension

One of the most important predictors of household extension, and a main factor explaining differences across groups, is socioeconomic status. As many researchers have shown, family extension tends to be negatively associated with such factors as household/personal income and educational attainment (Angel and Tienda 1982, Blank and Torrecilha 1998, Glick, Bean, and Van Hook 1997, Van Hook and Glick 2007). Economic theories of extension center on the idea that while individuals generally have a preference for independent living, the economies of scale and resulting reduction of living expenses that accrue to shared goods, such as housing, provides a powerful incentive for doubling up for low income groups (Browning, Chiappori, and Lewbel 2013, Deaton and Christina Paxson 1998, Deaton and Muellbauer 1980, Kakwani and Son 2005).

Thus, one reason that immigrants are more likely to live in extended family households is that many national origin groups average lower wages and education levels than the native born. Coupled with limited access to public assistance, extended living arrangements are an important way that immigrant households mitigate the consequences of economic insecurity and poverty (Glick and Van Hook 2002, Leach

2010). Moreover, because individuals may trade services within households, adults with limited labor market opportunities may form extended households in an attempt to exchange shelter and basic living costs for services such as the care of children and the elderly (Angel and Tienda 1982, Blank and Torrecilha 1998).

3.2.2. Life cycle theories of extension

Another theoretical perspective on household extension focuses on the link between living arrangements and the life-course. According to this perspective, the incentives for extension vary at different life stages; both biological age and specific events such as marriage, childbirth, and old age disability shape the incentives and opportunities for household extension (Blank and Torrecilha 1998, Moen and Wethington 1992).

When applied to immigrant-native differences in household formation, the argument is two-fold. First, immigrants are disproportionately young and average higher fertility than the native born. As such, they are more likely to form extended households because these life-cycle characteristics favor extension. Second, immigration is itself an important life-cycle event that can shape extension in and of itself. The period immediately following migration is often characterized by a high degree of uncertainty and insecurity, while new entrants into the society search for jobs and housing and become acclimated to their new environments. Thus, in the early stage of migration, extended household formation may emanate not only from low income, but also from the need to pool resources to recover from the insecurity and disrupted social bonds inherent in the migration experience. Later, as migrants gain experience and greater knowledge in United States, they may be better able to convert their economic resources into a nuclear household structure.

Empirical evidence supports most of these theoretical predictions. Studies of

Mexican immigrants show that recent migrants are more likely to live in extended households than those who have resided in the United States for longer periods of time (Glick, Bean, and Van Hook 1997, Van Hook and Glick 2007). Similarly, in analyses with cross-sectional data, being married is often negatively associated with a high likelihood of living in extended households (Blank and Torrecilha 1998). A strong positive correlation is also found between the presence of children under six years of age and extension (Blank and Torrecilha 1998). Further empirical evidence shows that the likelihood of extension varies by age among immigrants, and age composition helps explain much of the difference in extension between Mexican and Asian immigrants (Glick and Van Hook 2002).

3.2.3. Cultural theories of extension

While economic and life-cycle characteristics explain a large share of the higher propensity for household extension among immigrants, gaps with the native population remain even after these factors are accounted for. Cultural theories of extension rely on two main arguments to explain these remaining differences: taste formation (that the preference for nuclear living arrangements is lower in many developing regions) and portability of culture (that migrant household patterns in the United States represent a continuation of cultural practices from countries of origin). Given the absence of measures of culture and preferences, this framework is generally based on residual analysis, whereby the cultural explanation is derived from variation unexplained by socioeconomic and demographic variables (Angel and Tienda 1982, Blank and Torrecilha 1998, Van Hook and Glick 2007).

One obvious problem with attributing unexplained differences in household formation, that are often sizeable, to culture is that other omitted factors, often related to data availability, also contribute to the disparities (Blank and Torrecilha 1998, Van Hook and Glick 2007). Using a direct approach to measuring culture, Van Hook and Glick (2007) compare household structure in migrant sending regions in Mexico to household structure among Mexican Americans in the United States. They find a much smaller contribution of culture to Mexican American living arrangements than previous studies using a residual approach. To the extent that culture can evolve over time, one should expect changes in immigrant preferences as duration of residence in the United States increases. This has led many researchers to examine the link between acculturation and household extension. Findings tend to be sensitive to the measure of acculturation used. Using English proficiency as a proxy for assimilation, Blank and Torrecilha (1998) find no statistically significant relationship between culture and the likelihood of living in extended households (Blank and Torrecilha 1998). On the other hand, years since immigration are found to explain some variation in household composition, as the likelihood of extension tends to decrease as immigrants stay longer in the United States (Blank and Torrecilha 1998, Leach 2010, Van Hook and Glick 2007). Unfortunately, the interpretation of these results is complicated by the association of these variables with other socio-demographic and life-cycle characteristics, such as education, age of arrival to the United States, and so on.

3.2.4. New directions in research on extension among immigrants

Previous studies on household extension share some common shortcomings. First, they fail to consider the whole range of the immigrant experience. Most studies on extension have focused on Mexican and, to a lesser extent, Asian and other Latin American immigrants. Given the tremendous heterogeneity of immigrants from different sending regions, particularly with respect to the human capital characteristics associated with extension (Borjas 1987, 1994, Chiswick 1978, Elo et al. 2015, Feliciano 2005, Kritz and Gurak 2015, Lucas 2015), this is problematic. While a handful of studies compare multiple sending regions (see Glick, Bean, and Van Hook 1997), they tend to focus on earlier time periods. However, rapid social change abroad and the continued influx of newcomers to the United States continuously alter the composition of all immigrant groups, and necessitate reexamination of even familiar patterns. Second, most studies of immigrant living arrangements do not distinguish between the various types of extension. However, it should be emphasized that family extension encompasses three possible types of living arrangements: horizontal (with same-generation members), vertical (with multi-generation members), or mixed (horizontal and vertical) extension. This distinction has been shown to be important for Mexican immigrants (Van Hook and Glick 2007), and is in need of examination in a comparative perspective.

Finally, previous research often fails to account for the role of local context in structuring patterns of household extension. Because immigrant populations are more highly concentrated geographically than native-born groups, local housing market characteristics are likely to shape their living arrangements. The economic need to double up with family members is shaped by the income available to potential household members and also by the cost of housing that could be shared. Because low income is associated with family household extension, one would expect that high home values, ceteris paribus, would tend to increase the prevalence of family household extension. Other local conditions that suggest more difficult transitions into homeownership could also shape extension patterns. Thus, the greater concentration of immigrants in more expensive housing markets, and those marked by lower rates of home ownership, could contribute to the immigrant-native gap in extension.

Our paper addresses the limitations of previous studies in three principal ways.

First, we use recent census data to examine patterns of household extension among immigrants from Mexico, rest of America, the West Indies, South East Asia, rest of Asia, Africa, Oceania, and Europe. We are concerned not only with differences across groups in the prevalence of extension, but also in how the socio-demographic correlates of extension may differ across groups. Second, we distinguish between vertical, horizontal, and mixed extended households to explore differences across regions of origin, as well as how socio-demographic and life-cycle forces relate differently to these multiple forms of extension. And finally, we examine the role of the local context in shaping immigrant household extension, by region of origin and extension type, and assess the extent to which this dimension helps us to understand differences in patterns of household formation between the native born and the foreign born.

3.3. Data and Methods

We use data from the 2001 through 2013 waves of the American Community Survey (ACS), which is a representative sample of the non-institutionalized, civilian U.S. population. In the ACS annual release, each year pertains to approximately five percent of the U.S. population. The pooled data, containing both household and individual level information, are downloaded from the IPUMS website (Flood et al. 2015). The ACS contains extensive socioeconomic and life-cycle data along with a large sample size that allows the examination of each type of family extension among immigrants from diverse sending regions. It also contains geographic identifiers at the county level which allow for the inclusion of contextual characteristics.

We restrict the analytical sample to respondents aged 25 and above to be consistent with previous studies on household extension (Glick, Bean, and Van Hook 1997, Goldscheider and Bures 2003, Van Hook and Glick 2007). Our main objective is to investigate the differences in household extension between the native born and the foreign born. Immigrants are defined as those born outside of the United States (mainland and US territories). We compare immigrants from different origins to native-born non-Hispanic (NH) whites.[1] Because the NH white sample in the ACS is extremely large, we use a five percent subsample, which is drawn by stratifying on survey year, state of residence, gender, and age of the respondents. We further restrict the sample to respondents with non-negative (i.e., strictly positive or zero) income.[2] This yields an analytical sample of 3,382,048 respondents, 2,574,776 of whom are foreign-born.

3.3.1. Model specification

Family household extension, the main dependent variable under consideration, is defined using the concept of minimal household unit (MHU) (Glick, Bean, and Van Hook 1997, Glick and Van Hook 2002, Van Hook and Glick 2007). An MHU may consist of a household head, his/her spouse/partner, and any biological or foster children who are under 25 years old, unmarried, and have no children of their own. It may also consist of a single adult. A household is defined as an extended family household if it consists of at least two related MHUs. For the purpose of this analysis, we first determine whether a household contains an extended family member. We

[1]It is worth emphasizing that the native U.S. population is highly diverse, and using NH whites as the default reference against which immigrants are compared has rightly been criticized for conflating native with white. However, in the case of living arrangements, native Hispanic, Asian, and black populations are all more likely to live in extended households than their NH white peers (Angel and Tienda 1982). Indeed, the same economic and cultural theories of extension that have been applied to immigrants have also been used to explain ethno-racial variation in living arrangements within the native population. We therefore use NH whites as the comparison group not because they represent the native norm, but rather because they are the native group with the lowest level of extended living arrangements.

[2]Negative income is rarely reported; only 2,176 observations were excluded based on these criteria, representing 0.1% of the total sample (0.11% of Asian immigrants, 0.12% of NH native whites and West Indian immigrants, 0.08% of Latin American immigrants, 0.04% of African immigrants, and 0.05% of Canadian/European immigrants)

then use this information to define a household as extended when any of its members are extended kin.

Because a family can extend in three possible ways, we distinguish between horizontal, vertical, and mixed extension. Horizontally extended family households are those consisting of adults of the same generation. In other words, they refer to co-residence among siblings and other relatives (Glick, Bean, and Van Hook 1997, Glick and Van Hook 2002, Van Hook and Glick 2007). Vertically extended family households refer to those consisting of multiple generations. In other words, vertically extended family households include household heads who live with their parents or their adult children, grandparents co-residing with adult grandchildren, or of a combination of any of these living arrangements. Mixed extension applies to households that combine both horizontal and vertical structures (Glick, Bean, and Van Hook 1997, Glick and Van Hook 2002, Van Hook and Glick 2007).

Given the central role of immigrant origin to the analysis, we define binary indicators of major geographical sending regions, including Mexico, the West Indies (mostly consisting of the Caribbean), rest of America (South and Central America), Africa, South East Asia (China, India, Japan, and Korea), rest of Asia, Canada/Europe, and Oceania. We also tried alternative ways of grouping Asian and Latin American countries, but the results are substantively unaffected. Other independent variables in the analysis correspond to economic and life-cycle theories of extension. Measures of socioeconomic status include a binary variable that captures whether a respondent worked for pay during the week of reference of the survey, or whether the respondent worked at least 15 hours during the week on a family farm or business.[3] We also use total personal income, which consists of income from all sources pertaining to the 12 months preceding the survey. Income is modeled as a discrete variable, representing

[3] $http : //www.census.gov/people/laborforce/about/acs_employ.html$

income quintiles at the level of the pooled sample. We also include a measure of educational attainment, captured by four binary indicators for less than high school, high school, some college, and college graduate or higher. The ability and motivation to extend could also be influenced by housing tenure. We therefore also include a binary measure of whether or not the respondent owns their primary residence.

Other socio-demographic indicators include age (25-34, 35-44, 45-54, 55-64, or 65+), gender, and binary indicators of whether or not the respondent is married and whether or not they have a child under age five living with them. In models restricted to the foreign born we also include a number of immigrant characteristics, including a binary indicator of the ability to speak English well and years since immigration (0-9, 10-19, and 20 and or more).

Finally, we also examine the impact of local area housing characteristics on extension patterns. These include county-level median home value, homeownership rates. For local home values, the natural logarithm of median home value is computed at the county level and is ascribed to all respondents in the county. We also control for survey year to capture secular change in macroeconomic conditions.

3.3.2. Analytical strategy

We begin by extending the previous literature to all groups of immigrants and by distinguishing between horizontal, vertical, and mixed types of extension. A series of multinomial logit models are estimated to predict the likelihood of each type of extended family household with the nuclear household being the reference category. We compute White robust standard errors to account for clustering at the household level, which simultaneously accounts for clustering at the county level (Wooldridge 2010). We further account for the ACS sampling design by weighting all estimates by the personal weights provided in the ACS data files. All results are presented as relative risk ratios.

In Model 1, we control only for region of origin. This amounts to a bivariate analysis of the type of extension among immigrants by region of origin relative to NH white natives. Model 2 adds socioeconomic, demographic, and life cycle determinants of extension to Model 1 to investigate the extent to which the differences observed in Model 1 are attributable to differences in these socio-demographic characteristics across groups. Model 3 adds local county-level housing characteristics to Model 2 to examine what share of the unexplained residual difference among groups, often attributed to culture in previous studies, is related to the geographic concentration of immigrants in high-cost housing markets.

Finally, to further investigate whether the determinants of extension vary among the various immigrant groups we stratify the analyses by region of origin and limit the analyses to the foreign born. In addition to the explanatory variables included in Model 3 we also control for years since migration and English proficiency. All models also control for the survey year.

3.4. Results

Table 3.1 summarizes patterns of extension and socio-demographic characteristics by region of origin. The results clearly demonstrate the importance of distinguishing between different types of extension. While immigrants from all regions are significantly more likely than NH white natives to live in extended households, the overall extension rates in the first row mask considerable variation across regions in the relative importance of vertical and horizontal extension. For instance, emigrants from the rest of Asia (Asia excluding South East Asia and India) are more likely than other groups to extend overall, with an unadjusted prevalence of 42.5 percent, relative to

19.8 and 25.4 percent among NH white natives and Canadian/European immigrants, respectively (overall household extension among the other foreign-born groups ranges from 30.0 percent among Africans to 40.2 percent among Mexicans). However, while immigrants from the rest of Asia are the most likely to extend overall, this is primarily because they are far more likely than other groups to live in mixed extended households, with unadjusted prevalence of 12.6 percent for these immigrants relative to 3.8 and 5.2 percent for NH native whites and Canadians/Europeans, respectively. Also, the high prevalence of overall extension observed among Mexicans is mostly driven by their high propensity for horizontal extension. The unadjusted prevalence of horizontal extension is 18.4 percent among Mexicans compared to 7.2 and 8.2 percent among NH native white and Canadians/Europeans, respectively. Horizontal extension is also more common than vertical extension among immigrants from the rest of America (16.7 versus 14.7 percent), Africa (13.4 versus 10.1 percent), and West Indies (14.9 versus 14.5 percent), whereas the opposite is true for immigrants from South East Asia (10.3 versus 16.1 percent), the rest of Asia (13.5 versus 16.4 percent), Canada/Europe (8.2 versus 12.1 percent), and Oceania (7.2 versus 8.7 percent).

At least some of the large differences in living arrangements across groups are no doubt related to socio-demographic and life cycle characteristics, which also vary substantially across groups. Mexicans are younger, on average, than NH white natives, and South East Asians have high rates of marriage relative to other groups. All immigrants, with the exception of Canadians/Europeans, are more likely to have young children at home than native NH whites. But the largest differences across groups relate to socioeconomic factors. South East Asian, Oceanian, and Canadian/European immigrants average higher levels of both education and income than native NH whites. While Africans are also more likely to have college degrees than native NH whites, they nevertheless lag behind them in terms of income, though to a far lesser extent than their peers from the remaining regions. Overall, Mexicans average the lowest incomes and education levels of all groups considered. Africans average higher employment levels than other groups; and native whites are far more likely than others to be homeowners.

There are also marked differences across regional groups in their average immigration characteristics. Canadians/Europeans average the longest durations in the United States, followed by Mexicans. Africans, in contrast, are far more recently arrived than the other groups. Immigrants from Mexico are less likely to have a good command of English (43 percent relative to 86, 78, 75, 71, 60, 60, and 53 percent of immigrants from Oceania, Canada/Europe, Africa, West Indies, South East Asia, the rest of Asia, and the rest of America, respectively).

The potential importance of local context is also evident in Table 3.1. Native NH whites are not only more likely to own a home, they also live in counties with higher homeownership rates than their immigrant peers. This is in part because native NH whites tend to live in counties with lower average housing values. Specifically, the median home value in native NH whites counties of residence is 132,600, and 70 percent of area residents are homeowners. In contrast, the median home values in the counties in which South East Asian immigrants live is 182,000 and only 66 percent of area residents own their homes.

We next examine differences in extension across groups that remain after accounting for socio-demographic, life cycle, and contextual variation across groups. Table 3.2 presents relative risk ratios from multinomial logistic regression models of horizontal, vertical, and mixed extended family living arrangements. The first three columns of the table pertain to Model 1, which produces unadjusted relative risk ratios of each type of extension relative to the nuclear household structure, with NH white natives as the reference group. Once again we see that immigrants are far more likely than the native born to extend, with important differences by type of extension and region of origin. Results show the same patterns evident in Table 3.1: Mexican immigrants are 3.5 and 3.6 times more likely than native NH whites to live in horizontal and mixed extension households, respectively, but only 2.0 times more likely to live in vertically extended households. Once again, Asian (South East Asia and Rest of Asia), Canadian/European, and Oceanian immigrants are distinct from other groups in that differences in extension with native NH whites are larger for vertical than horizontal extension; for all other groups the opposite is true.

The next three columns, for Model 2, introduce controls for individual and household level characteristics. As in previous studies both life cycle and economic conditions shape extension patterns, though there are interesting differences across types of extension. For instance, older ages are associated with lower odds of living in horizontally extended households, but higher odds of living in vertically extended households. Married respondents are less likely than their non-married peers to extend, but this tendency is especially pronounced for horizontal extension. Having young children discourages horizontal extension but increases the likelihood of vertical extension. Likewise, there is a strong inverse relationship between income and horizontal extension, but the relationship is weaker for vertical extension; only the lowest income quintiles differ significantly from the vertical extension patterns of the top income group, and they are more likely to extend, consistent with economic theories of extension. Employment is also a significant predictor of a greater likelihood of horizontal but not vertical extension. Higher levels of education generally discourage extension, while homeownership facilitates it.

More importantly for our purposes, the inclusion of socio-demographic and life cycle controls in Model 2 also explains part of the immigrant-native disparities in household extension, although the results vary by sending region and type of extension. To facilitate the comparisons across groups and types of extension, Figures 3.1 and 3.2 graph the relative risk ratios of horizontal and vertical extension, respectively, across the 3 models shown in Table 3.2 (unadjusted Model 1, controlling for socio-demographic and life cycle characteristics Model 2, and controlling for socio-demographic, life cycle and contextual characteristics Model 3). Two intriguing patterns are evident: differences in the impact of controls on relatively more and less advantaged immigrant groups, and differences in the impact of controls on horizontal and vertical extension. Figure 3.1 shows that for the three least advantaged groups, Mexicans, immigrants from the rest of America, and West Indians, controlling for socio-demographic characteristics substantially reduces the gap in horizontal living arrangements with respect to native NH whites. For Asian (South East Asia and Rest of Asia), Canadian/European, and Oceanian immigrants, in contrast, controlling for their socio-demographic characteristics slightly increases their likelihood of horizontal extension compared to native NH whites. For African immigrants, the likelihood of horizontal extension is unchanged. A very different pattern is evident for vertical extension, presented in Figure 3.2. Overall socio-demographic and life cycle characteristics explain much less of the variation by region of origin, and controlling for these factors raises the gap in vertical extension between native NH whites and all immigrant groups, especially for Asians (South East Asia and rest of Asia).

Finally, Model 3 adds indicators of local housing context. Results indicate that net of individual and household level socio-demographic and life cycle characteristics local housing conditions also have an important impact on living arrangements. That is, both vertical and horizontal extension are more common in areas where housing costs are higher and access to home ownership is more limited. Moreover, accounting for the uneven geographic concentration of groups across the country has an important impact on native-immigrant extension differentials, as illustrated in Figures 3.1

and 3.2. Comparing the second and third bars across groups shows that accounting for immigrants tendency to live in more expensive and lower homeownership contexts reduces the native-immigrant differences in extension for all regions of origin and for both horizontal and vertical extension. As was the case with individual level controls, however, context seems to matter more for horizontal living arrangements than for vertical extension. There are also important differences across regions of origin, as the Canadian/European differentials appear to be less influenced by context than for the other immigrant groups for both types of extension.

The pronounced region-of-origin differentials in household extension evident in Tables 1 and 2 invite further examination. In addition to differences in the pattern and prevalence of household extension across regions of origin, there are also potential differences in the impact of explanatory variables on extension patterns across groups. To explore this possibility, Tables 3.3 and 3.4 present relative risk ratios for horizontal, vertical, and mixed extension relative to nuclear living arrangements, separately for our height regions of origin from the fully adjusted model which includes years since arrival in the U.S. and English fluency in addition to controls used in Model 3, Table 3.2.

Overall, most socio-demographic, life cycle and contextual characteristics have roughly comparable impacts on extension across groups. However, there are some important exceptions. The biggest differences across groups are evident in the impact of income and duration of U.S. residence on living arrangements. To better illustrate these differences, Figures 3.3 and 3.4 graph the fully adjusted risk ratios of horizontal and vertical extension, respectively, relative to nuclear living arrangements, by income for each region of origin. Figure 3.5, in turn, graphs the fully adjusted risk ratios of extension by duration of U.S. residence and by region of origin. Both figures are based on the risk ratios presented in Tables 3.3 and 3.4

Beginning with Figures 3.3 and 3.4, results show considerable variation in the association between income and extension both by region of origin and type of extension. While Tables 3.3 and 3.4 show that virtually all of the lower income quintiles are statistically more likely than those in the top quintile to live in horizontally extended households, regardless of region of origin, Figure 3.3 shows that in substantive terms the income gradient is more muted for the relatively advantaged immigrant groups, namely South East Asians, Canadians/Europeans, and Oceanians. The gradient is far steeper among the more disadvantaged groups (Mexicans, West Indians, and immigrants from the rest of America) and immigrants from Africa. For these groups, the income gradient for horizontal extension is nonlinear, with the lowest income quintiles far more likely than others to horizontally extend.

With respect to vertical extension, in contrast, both the impact of income and differences across regions are far more modest. Figure 3.4 shows that for all immigrants, the likelihood of vertical extension decreases with income, although most differences are not statistically significant for either group.

Finally, Figure 3.5 illustrates the fully adjusted relative risk ratios for both horizontal and vertical extension relative to nuclear living arrangements by region of origin and duration of U.S. residence, with fewer than 10 years in the United States as the reference category. For all groups the relative risk of horizontal extension decreases with longer duration of U.S. residence. For vertical extension, on the other hand, duration in the United States has little effect among Mexicans. For most other migrant groups, the likelihood of extension decreases with longer duration in the United States, sometimes after a peak in the decade following the time of arrival.

3.5. Conclusion

The foreign born represent a rapidly growing share of the US population. According to the 2010 ACS, the foreign born made up 13 percent of the total U.S. population, amounting to about 40 million people (Grieco et al. 2012). Not only is the share of the foreign born of the US population growing, the foreign born are becoming increasingly heterogeneous. For example, among all foreign-born migrants, those born in Africa have been growing at the fastest rate in recent years (Elo et al. 2015) with Asians projected to become the largest immigrant group surpassing Hispanics by 2050 (Lopez, Passel, and Rohal 2015). And yet, studies that have investigated co-residence among the foreign-born population have focused primarily on immigrants from Mexico (Lopez, Passel, and Rohal 2015) and on those from a few Asian countries (Lopez, Passel, and Rohal 2015). Although some previous studies include more expansive comparisons (Glick, Bean, and Van Hook 1997), the continuous evolution of immigrant composition requires frequent updating to assess contemporary patterns. In addition, most studies fail to distinguish between different types of extension or examine the role of local contextual conditions in shaping differences in living arrangements across immigrant groups. The large sample size that results from pooling the 2001 through 2013 waves of the ACS makes it possible to investigate the relevance of the leading theories of household extension among multiple immigrant groups that are representative of the foreign-born population. In addition, we extend prior research by examining the predictors of both horizontal and vertical extension, and by investigating whether local housing market conditions influence immigrants household extension patterns and their divergence from those of native NH whites.

Our findings lend mixed support to the findings from previous research on extension. While we show that immigrants have a higher propensity for all types of extension than native-born NH whites, as in previous studies (Blank and Torrecilha 1998, Glick 1999, Glick, Bean, and Van Hook 1997, Glick and Van Hook 2002, Kamo 2000, Van Hook and Glick 2007), we also find substantial variation among different types of extension and by region of origin.

Our findings strongly suggest that horizontal and vertical extension are driven by different forces. Economic and life-cycle theories of extension do a far better job of explaining horizontal than vertical extension, suggesting that greater conceptual and empirical clarity is needed when evaluating immigrants engagement in household extension.

Economic and life-cycle characteristics also tend to hold more sway among immigrants from less advantaged origins. Not only do socio-demographic and life cycle characteristics explain a larger share of the immigrant-native differences in horizontal extension for Mexicans, West Indians, and immigrants from the rest of America than for South East Asians, Canadians/Europeans, and Oceanians relative to native NH whites, the income gradient for horizontal extension is steeper among these less advantaged region-of-origin groups. The income gradient for horizontal extension among Africans, who on average have higher educational attainment but lower incomes than native NH whites, is also steeper than among the more advantaged immigrant groups. In contrast to horizontal extension, socio-demographic and life cycle characteristics do little to explain patterns of vertical household extension among the various immigrant groups relative to the native NH whites. Furthermore, compared to horizontal extension the income gradient is relatively flat among all origin groups when it comes to vertical household extension.

Our findings also highlight the importance of geographic distribution of immigrants and the need to take into account the role of housing conditions in studies of household extension. Immigrants are disproportionately concentrated in housing markets characterized by high costs and low rates of homeownership, both of which tend to encourage extended living. Accounting for differences in local context across groups explains up to 10 percent of the difference in extension between the foreign-born subgroups and native-born NH whites. These results provide new evidence against the interpretation that the unexplained differences among the foreign born and the native born are due simply to cultural differences that shape immigrant living arrangements. In all probability, including a richer set of community characteristics could go even further towards explaining immigrant-native differences in household extension.

This study has a number of limitations. First, while we identify an important source of the unexplained variation in household extension across region-of-origin groups (namely local context), we are unable to explain all of the variation in household extension patterns among the foreign-born subgroups relative to native NH whites. In particular, the lack of information on legal status could influence our findings, as undocumented immigrants suffer a number of disadvantages (Flippen 2012, Glick, Bean, and Van Hook 1997) that could also shape extension patterns. There could also be additional county level or local area characteristics that influence extension. Finally, the variable regarding the duration since immigration is computed by subtracting the year of arrival from survey year. The shortcoming of this commonly used approach is that it may not accurately capture duration since immigration, especially among immigrants who cross international borders multiple times (Massey 1987).

Immigrant living arrangements have long captured the attention of researchers both as a sign of economic marginality and as a marker of cultural distinctiveness from the native mainstream. The strong link between economic hardship and horizontal extension, which is both more prevalent and has a stronger income gradient among more disadvantaged immigrant groups, indeed warrants concern. These patterns suggest that household-level analyses underestimate the financial vulnerability of low-income families from Mexico, the West Indies, the rest of America, and Africa. The steady reductions in both horizontal and vertical extensions with longer durations in the United States, in contrast, support the argument that extension is less an enduring cultural artifact than a temporary stage in the process of immigrant adaptation. The fact that vertical extension is both less related to socio-demographic characteristics and less variable across region-of-origin groups implies that this particular living arrangement could have more to do with the needs and conditions of elder generations than with those of working-aged adults. In addition, the immigrants residential context matters for patterns of household extension, an area that has not received much prior attention. The cost of living, which we measured by the cost of housing in the county of residence, is important. Future studies of household living arrangements should thus consider not only individual- and household-level characteristics but also take into account residential context.

3.6. References

Angel, R., and M. Tienda. 1982. "Determinants of Extended Household Structure: Cultural Pattern or Economic Need?" American Journal of Sociology 87 (6):1360-1383.

Blank, S., and R. S. Torrecilha. 1998. "Understanding the Living Arrangements of Latino Immigrants: A Life Course Approach." International Migration Review 32 (1):3-19.

Borjas, G. 1987. "Self-Selection and the Earnings of Immigrants." American Economic Review 77 (4):531-553.

Borjas, G. 1994. "The Economics of Immigration." Journal of Economic Literature XXXII:pp. 16671717.

Browning, M., A. Chiappori, and A. Lewbel. 2013. "Estimating Consumption Economies of Scale, Adult Equivalent Scales, and household Bargaining Power." The Review of Economic Studies 80 (4):1267-1303.

Camarota, S. A. 2012. Immigrants in the United States, 2010: A Profile of America's Foreign-Born Population. Washington, DC: Center for immigration Studies.

Chiswick, B. R. 1978. "The Effect of Americanization on the Earnings of Foreign-Born Men." The Journal of Political Economy 86 (5):897-922.

Deaton, A., and C. Christina Paxson. 1998. "Economies of scale, Household Size, and the Demand for food." The Journal of Political Economy Vol. 106:897-930.

Deaton, A., and J. Muellbauer. 1980. Economics and Consumer Behavior: Cambridge,UK: Cambridge University Press.

Elo, Irma T, Elizabeth Frankenberg, Romeo Gansey, and Duncan Thomas. 2015. "Africans in the American Labor Market." Demography 52 (5):1513-1542.

Feliciano, Cynthia. 2005. "Educational selectivity in US immigration: How do immigrants compare to those left behind?" Demography 42 (1):131-152.

Flippen, Chenoa A. 2012. "Laboring underground: The employment patterns of Hispanic immigrant men in Durham, NC." Social Problems 59 (1):21.

Flood, S., K. Miriam, S. Ruggles, and J. R. Warren. 2015. "Integrated Public Use Microdata Series, Current Population Survey, 2015." Version 4.0. [Machine-readable database]. Minneapolis: University of Minnesota.

Glick, J. E. 1999. "Economic Support from and to Extended Kin: A Comparison of Mexican Americans and Mexican Immigrants." International Migration Review 33 (3):745-765.

Glick, J. E., F. D. Bean, and J. V. W. Van Hook. 1997. "Immigration and Changing Patterns of Extended Family Household Structure in the United States: 1970-1990." Journal of Marriage and Family 59 (1):pp. 177-191.

Glick, J. E., and J. V. W. Van Hook. 2002. "Parents' Coresidence with Adult Children: Can Immigration Explain Racial and Ethnic Variation?" Journal of Marriage and Family 64 (1):240-253.

Goldscheider, F. K., and R. M. Bures. 2003. "The Racial Crossover in Family Complexity in the United States." Demography 40 (3):569-587.

Grieco, Elizabeth M. , Yesenia D. Acosta, G. Patricia de la Cruz, Christine Gambino, Thomas Gryn, Luke J. Larsen, Edward N. Trevelyan, and Nathan P. Walters. 2012. "The Foreign-Born Population in the United States: 2010." US Census Bureau.

Jasso, Guillermina. 2011. "Migration and stratification." Social Science Research 40 (5):1292-1336.

Jasso, Guillermina, Douglas S. Massey, Mark R. Rosenzweig, and James P. Smith. 2004. Immigrant health: selectivity and acculturation. IFS Working Papers, Institute for Fiscal Studies (IFS).

Kakwani, N. , and H. Son. 2005. "Economies of Scale in Household Consumption:

With Application to Autralia." Australian Economic Papers 44 (2):134-148.

Kamo, Y. 2000. "Racial and ethnic differences in extended family households." Sociological Perspectives 43 (2):211-229.

Kritz, Mary M, and Douglas T Gurak. 2015. "US Immigrants in Dispersed and Traditional Settlements: National Origin Heterogeneity." International Migration Review 49 (1):106-141.

Leach, M. A. 2010. "Household Living Arrangements and Economic Resources among Mexican Immigrant Families with Children." Population Research Institute The Pennsylvania State University.

Lopez, M. H., J. Passel, and M. Rohal. 2015. "Modern Immigration Wave Brings 59 Million to U.S., Driving Population Growth and Change Through 2065: Views of Immigrations Impact on U.S. Society Mixed." Pew Research Center.

Lucas, Robert E. B. 2015. "African Migration." Handbook of the Economics of International Migration 1:1445-1596.

Massey, D. 1987. "Understanding Mexican Migration to the United States." American Journal of Sociology 92:1372-1402.

Moen, P., and E. Wethington. 1992. "The Concept of Family Adaptive Strategies." Annual Review of Sociology 18:233-251.

Van Hook, J. V. W., and J. E. Glick. 2007. "Immigration and Living Arrangements: Moving beyond Economic Need versus Acculturation." Demography 44 (2):225-249.

Wooldridge, Jeffrey M. 2010. Econometric analysis of cross section and panel data: Cambridge, MA: MIT press.

3.7. Tables

Table 3.1: Descriptive statistics for non-Hispanic native-born and foreign-born ages 25+: 50 states and the District of Columbia, ACS 2001-2013 (percentage unless otherwise noted)

Characteristics (N/1,000)	All (3,382)	Mexico (598)	W. Indies (141)	Rest Am. (400)	Africa (84)	S. E. Asia (392)	Rest Asia (372)	Europe (571)	Oceania (14)	US-b. (807)
Extension (all types)	33.36	41.23	38.05	41.03	30.01	34.95	42.52	25.45	32.71	19.80
Horizontal exten.	12.57	18.48	14.88	16.67	13.40	10.26	13.54	8.19	11.22	7.23
Vertical exten.	12.84	12.54	14.48	14.66	10.84	16.08	16.41	12.06	12.49	8.75
Mixed exten.	7.95	10.21	8.69	9.70	5.78	8.62	12.57	5.20	9.00	3.82
Gender										
[Female]										
Male	53.26	63.17	45.62	53.52	58.08	53.46	49.24	48.18	52.73	49.95
Mean age	47.80	42.12	47.43	46.11	43.15	46.62	47.80	53.96	46.04	51.54
[SE]	[0.01]	[0.02]	[0.05]	[0.03]	[0.05]	[0.03]	[0.03]	[0.03]	[0.16]	[0.02]
Age										
25-34 years	22.89	33.35	19.87	24.99	27.43	23.86	20.20	14.86	24.81	17.75
35-44 years	25.42	31.29	26.80	27.49	31.79	26.83	26.62	19.44	28.11	19.84
45-54 years	21.33	18.96	24.90	21.89	23.97	22.22	23.03	19.99	21.28	21.77
55-64 years	14.03	9.09	15.21	12.47	10.71	13.62	15.74	16.45	13.71	17.82
65 years+	16.33	7.30	13.22	13.16	6.10	13.47	14.41	29.25	12.09	22.82
Marital status										
Not married	36.32	35.24	48.63	42.60	37.69	24.73	32.10	35.63	33.37	39.30
Curr. married	63.68	64.76	51.37	57.40	62.31	75.27	67.90	64.37	66.63	60.70
Child $<$ 5										
No	96.92	95.34	97.57	97.44	93.97	97.16	96.78	97.87	95.86	97.69
Yes	3.08	4.66	2.43	2.56	6.03	2.84	3.22	2.13	4.14	2.31
Mean income (x1,000)	18.26	11.21	14.83	14.68	20.30	25.19	19.67	22.77	25.65	20.30
[SE]	[0.02]	[0.02]	[0.05]	[0.03]	[0.11]	[0.06]	[0.05]	[0.05]	[0.34]	[0.03]
Income										
bottom quintile	19.75	23.07	21.18	21.48	16.75	18.53	20.47	17.90	16.59	17.05
2nd quintile	20.91	30.81	21.18	24.93	18.04	14.23	16.68	17.25	14.91	17.23
3rd quintile	20.90	25.89	23.16	23.42	21.90	15.06	19.23	18.12	18.18	19.68
4th quintile	19.97	14.84	21.93	18.57	22.12	19.23	21.46	21.14	21.54	24.10
Top quintile	18.47	5.39	12.55	11.61	21.19	32.95	22.16	25.59	28.78	21.94
Educational attainment										
$<$ high school	22.99	55.13	21.16	27.44	6.86	9.55	14.42	11.57	9.64	10.51
High school	31.10	30.20	37.87	34.61	25.44	18.70	25.51	31.81	32.21	38.10
Some college	17.17	9.03	21.68	17.42	23.10	12.62	20.09	19.93	21.64	22.28
College +	28.74	5.64	19.28	20.53	44.60	59.13	39.98	36.69	36.50	29.11
Is employed										
No	27.49	20.16	25.68	23.19	17.40	23.21	26.62	37.62	23.26	34.98
Yes	72.51	79.84	74.32	76.81	82.60	76.79	73.38	62.38	76.74	65.02
Owns home										
No	38.13	50.81	49.45	48.53	52.65	37.16	33.07	27.80	40.46	24.40
Yes	61.87	49.19	50.55	51.47	47.35	62.84	66.93	72.20	59.54	75.60
Speaks English only, or very well										
No	-	56.09	28.13	46.79	24.52	39.59	39.96	21.07	13.95	-
Yes	-	43.91	71.87	53.21	75.48	60.41	60.04	78.93	86.05	-
Mean years since immigration		19.87	21.67	20.46	15.77	19.45	21.19	30.77	20.74	
[SE]	-	[0.02]	[0.04]	[0.03]	[0.05]	[0.03]	[0.02]	[0.03]	[0.16]	-
Years since immigration										
0-9	-	22.72	17.01	24.87	37.56	27.05	18.25	15.99	30.00	-
10-19	-	31.57	29.42	27.92	30.91	28.10	27.26	17.91	22.96	-
20+	-	45.71	53.57	47.21	31.54	44.86	54.49	66.09	47.04	-
County level proportion of homeownership		66.47	62.81	65.78	66.20	66.17	66.24	68.26	66.22	70.61
Median county level home value (x1,000)	152.90	152.90	167.20	167.20	167.20	182.40	179.20	152.90	171.70	132.60

Frequencies (/1,000) in parentheses
Note: Sample characteristics are based on weighted data. The number of cases is unweighted
Source: 5% PUMS file of ACS waves 2001-2013

Table 3.2: Relative risk ratios from multinomial logistic regression models of extended family living arrangements among non-Hispanic native-born and foreign-born ages 25+, by type of extension: 50 states and the District of Columbia, ACS 2001-2013 (N=3,382,048)

Characteristics	Model 1			Model 2			Model 3		
	Horizon.	Vertical	Mixed	Horizon.	Vertical	Mixed	Horizon.	Vertical	Mixed
Sending regions									
[NHW US-born]									
Mexico	3.52***	1.97***	3.68***	2.75***	2.22***	3.46***	2.67***	2.16***	3.29***
West indies	2.67***	2.14***	2.95***	2.41***	2.30***	3.02***	2.03***	1.99***	2.31***
Rest America	3.17***	2.30***	3.48***	2.79***	2.53***	3.53***	2.52***	2.32***	3.05***
Africa	2.15***	1.43***	1.75***	2.15***	1.81***	2.09***	1.97***	1.68***	1.81***
South East Asia	1.76***	2.28***	2.80***	1.97***	2.73***	3.38***	1.76***	2.46***	2.84***
Rest Asia	2.63***	2.63***	4.62***	2.73***	2.93***	5.07***	2.47***	2.68***	4.37***
Europe/Canada	1.22***	1.48***	1.46***	1.30***	1.46***	1.52***	1.21***	1.38***	1.38***
Oceania	1.85***	1.70***	2.81***	1.91***	1.98***	3.19***	1.72***	1.81***	2.74***
Gender									
[Female]									
Male				1.17***	0.90***	1.07***	1.18***	0.91***	1.08***
Age									
[25-34 years]									
35-44 years				0.87***	1.10***	0.82***	0.86***	1.09***	0.81***
45-54 years				0.80***	1.23***	0.84***	0.79***	1.21***	0.82***
55-64 years				0.79***	1.36***	0.92***	0.77***	1.33***	0.88***
65 years +				0.70***	1.81***	0.99	0.68***	1.76***	0.94***
Marital status									
[Not married]									
Married				0.65***	0.79***	0.74***	0.66***	0.80***	0.75***
Child <5									
[No]									
Yes				0.76***	1.10***	0.91***	0.77***	1.11***	0.93***
Income quintile									
[Top quintile]									
bottom quintile				1.35***	1.23***	1.56***	1.42***	1.28***	1.68***
2nd quintile				1.34***	1.05***	1.35***	1.41***	1.09***	1.45***
3rd quintile				1.21***	1.00	1.23***	1.26***	1.04***	1.30***
4th quintile				1.10***	1.01	1.14***	1.12***	1.03***	1.18***
Educational attainment									
[high school]									
< high school				1.16***	1.11***	1.13***	1.15***	1.10***	1.12***
some college				0.85***	0.95***	0.92***	0.85***	0.95***	0.92***
college +				0.78***	0.88***	0.78***	0.78***	0.87***	0.77***
Owns home									
[No]									
Yes				1.09***	1.29***	1.50***	1.17***	1.38***	1.69***
Is employed									
[No]									
Yes				1.13***	0.98***	1.06***	1.14***	0.99**	1.07***
County level homeownership							1.66***	1.56***	1.55***
ln(County level median home value)							1.40***	1.35***	1.60***
Control for survey year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Source: 5% PUMS file of ACS waves 2001-2013
Note: Estimates are weighted. Reference group for each variable is given in brackets.
*** p<0.01, ** p<0.05, * p<0.1

Table 3.3: Relative risk ratios from a series of multinomial logistic regression models of extended family living arrangements among foreign-born ages 25+, by type of extension: 50 states and the District of Columbia, ACS 2001-2013

Characteristics		Mexico			West Indies			Rest America			Africa		
		Horizon.	Vertical	Mixed	Horizon.	Vertical	Mixed	Horizon.	Vertical	Mixed	Horizon.	Vertical	Mixed
Gender													
	[Female]												
	Male	1.28***	0.88***	1.10***	1.19***	0.84***	1.06*	1.21***	0.85***	1.06***	1.03	0.81***	0.94
Age													
	[25-34 years]												
	35-44 years	0.90***	1.06***	0.78***	0.93**	1.00	0.84***	0.88***	1.12***	0.85***	0.88***	1.23***	0.91
	45-54 years	0.83***	1.40***	0.91***	0.89***	1.15***	0.89***	0.84***	1.31***	0.94**	0.81***	1.34***	0.91
	55-64 years	0.78***	1.71***	1.11***	0.96	1.32***	0.99	0.82***	1.61***	1.02	0.78***	1.54***	0.91
	65 years+	0.70***	2.53***	1.26***	0.93	2.02***	1.30***	0.78***	2.15***	1.16***	0.74***	2.52***	1.24**
Marital status													
	[Not married]												
	Married	0.62***	0.85***	0.75***	0.64***	0.82***	0.69***	0.63***	0.79***	0.72***	0.66***	0.89***	0.68***
Child <5													
	[No]												
	Yes	0.71***	1.02	0.91***	0.74***	1.14*	0.71***	0.81***	1.01	0.79***	0.60***	1.12*	0.74***
Income quintile													
	[Top quintile]												
	bottom quintile	1.35***	1.10***	1.46***	1.40***	1.26***	1.66***	1.49***	1.27***	1.78***	1.33***	1.11*	1.24***
	2nd quintile	1.35***	1.01	1.40***	1.27***	1.03	1.33***	1.47***	1.11***	1.56***	1.20***	0.95	1.05
	3rd quintile	1.20***	1.00	1.23***	1.24***	1.02	1.24***	1.32***	1.09***	1.40***	1.14***	0.89**	1.06
	4th quintile	1.09***	1.04	1.16***	1.16***	1.00	1.05	1.18***	1.08***	1.25***	1.07	0.96	0.96
Educational attainment													
	[high school]												
	< high school	1.10***	1.02	1.07***	1.03	1.05	1.05	1.31***	0.98	1.09***	0.90	1.13*	1.25***
	some college	0.86***	0.99	0.98	0.90***	0.97	0.91**	0.80***	0.92***	0.85***	0.77***	0.99	0.94
	college +	0.76***	0.89***	0.78***	0.82***	0.89***	0.90**	0.71***	0.82***	0.75***	0.71***	0.91**	0.82***
Owns home													
	[No]												
	Yes	1.15***	1.25***	1.63***	1.55***	1.47***	1.91***	1.31***	1.43***	1.88***	1.53***	1.43***	1.77***
Is employed													
	[No]												
	Yes	1.13***	0.96***	1.04**	1.12***	1.03	1.14***	1.21***	0.99	1.12***	1.23***	1.00	1.01
Speaks English only, or very well													
	[No]												
	Yes	1.03***	1.08***	1.09***	0.82***	0.90***	0.87***	0.87***	1.00	0.97	0.87***	0.90***	0.97
Years since immigration													
	[0-9]												
	10-19	0.73***	1.03*	0.88***	0.73***	0.93**	0.80***	0.77***	0.94***	0.95**	0.79***	1.14***	1.05
	20+	0.57***	0.99	0.74***	0.61***	0.77***	0.65***	0.64***	0.83***	0.77***	0.72***	0.97	0.91
County level homeownership		1.31***	1.10	1.27**	1.91***	1.47**	2.96***	1.98***	1.05	0.82*	1.34	2.79***	1.07
ln(County level median home value)		1.69***	1.61***	2.11***	1.16***	1.05**	1.23***	1.30***	1.15***	1.31***	1.33***	1.25***	1.29***
Control for survey year		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Source: 5% PUMS file of ACS waves 2001-2013
Note: Estimates are weighted. Reference group for each variable is given in brackets.
*** p<0.01, ** p<0.05, * p<0.1

Table 3.4: Relative risk ratios from a series of multinomial logistic regression models of extended family living arrangements among foreign-born ages 25+, by type of extension: 50 states and the District of Columbia, ACS 2001-2013 (Continued)

Characteristics		S. E. Asia			Rest Asia			Europe			Oceania		
		Horizon.	Vertical	Mixed	Horizon.	Vertical	Mixed	Horizon.	Vertical	Mixed	Horizon.	Vertical	Mixed
Gender													
	[Female]												
	Male	1.14***	1.02	1.20***	1.06***	0.92***	1.05***	1.08***	0.94***	1.08***	0.96	0.98	1.11
Age													
	[25-34 years]												
	35-44 years	0.95**	1.09***	0.74***	1.03	1.15***	0.84***	0.94**	1.03	0.92***	0.95	0.95	0.75**
	45-54 years	0.96	1.15***	0.75***	1.06**	1.21***	0.82***	0.96	1.15***	0.91***	0.91	1.05	0.72**
	55-64 years	0.95*	1.27***	0.88***	1.11***	1.36***	0.95**	0.95**	1.23***	0.96	0.86	1.03	0.72**
	65 years+	0.79***	1.86***	0.95	0.98	2.32***	1.29***	0.84***	1.36***	0.88***	0.87	1.35**	0.86
Marital status													
	[Not married]												
	Married	0.71***	0.83***	0.77***	0.56***	0.78***	0.63***	0.73***	0.70***	0.78***	0.85*	1.09	0.81**
Child <5													
	[No]												
	Yes	0.96	1.27***	1.20***	0.78***	1.35***	1.00	0.89**	1.01	0.89*	0.88	1.08	1.16
Income quintile													
	[Top quintile]												
	bottom quintile	1.32***	1.36***	1.77***	1.34***	1.30***	1.77***	1.20***	1.25***	1.41***	1.20	1.46***	2.54***
	2nd quintile	1.33***	1.08***	1.50***	1.27***	0.98	1.42***	1.13***	1.18***	1.25***	1.57***	1.25*	1.80***
	3rd quintile	1.30***	1.05**	1.37***	1.32***	0.99	1.37***	1.11***	1.05**	1.17***	1.39**	1.29**	1.77***
	4th quintile	1.19***	1.01	1.29***	1.22***	1.00	1.26***	1.08***	1.04**	1.11***	1.31**	1.02	1.59***
Educational attainment													
	[high school]												
	< high school	1.10***	1.28***	1.25***	1.01	1.25***	1.22***	1.07***	1.21***	1.09***	1.03	1.36***	1.69***
	some college	0.82***	0.92***	0.87***	0.93***	1.00	0.97	0.95**	0.93***	0.95**	0.94	0.98	0.86
	college +	0.66***	0.79***	0.63***	0.81***	0.89***	0.79***	0.88***	0.89***	0.83***	0.59***	0.69***	0.47***
Owns home													
	[No]												
	Yes	1.40***	1.73***	2.28***	1.50***	1.65***	2.23***	1.15***	1.42***	1.40***	1.02	1.20**	1.57***
Is employed													
	[No]												
	Yes	1.08***	0.97*	1.05**	1.03	0.98	1.02	0.97	0.97*	0.97	0.92	1.05	1.17
Speaks English only or very well													
	[No]												
	Yes	0.91***	0.89***	0.88***	0.84***	0.88***	0.79***	0.83***	0.75***	0.77***	0.50***	0.55***	0.53***
Years since immigration													
	[0-9]												
	10-19	0.96*	1.18***	1.17***	0.82***	1.03	1.02	0.97	1.00	0.99	0.99	1.15	1.19
	20+	0.79***	0.81***	0.76***	0.68***	0.77***	0.65***	0.84***	0.74***	0.73***	0.79**	1.00	0.96
County level homeownership		1.05	1.95***	1.61***	0.84	1.06	0.82*	1.16	1.88***	1.36**	9.42***	4.55**	11.83***
ln(County level median) home value		1.37***	1.37***	1.64***	1.50***	1.45***	1.78***	1.30***	1.35***	1.48***	1.67***	1.48***	2.01***
Control for survey year		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Source: 5% PUMS file of ACS waves 2001-2013
Note: Estimates are weighted. Reference group for each variable is given in brackets.
*** p<0.01, ** p<0.05, * p<0.1

3.8. Figures

Figure 3.1: Relative Risk Ratios of Horizontal Extension by Sending Region

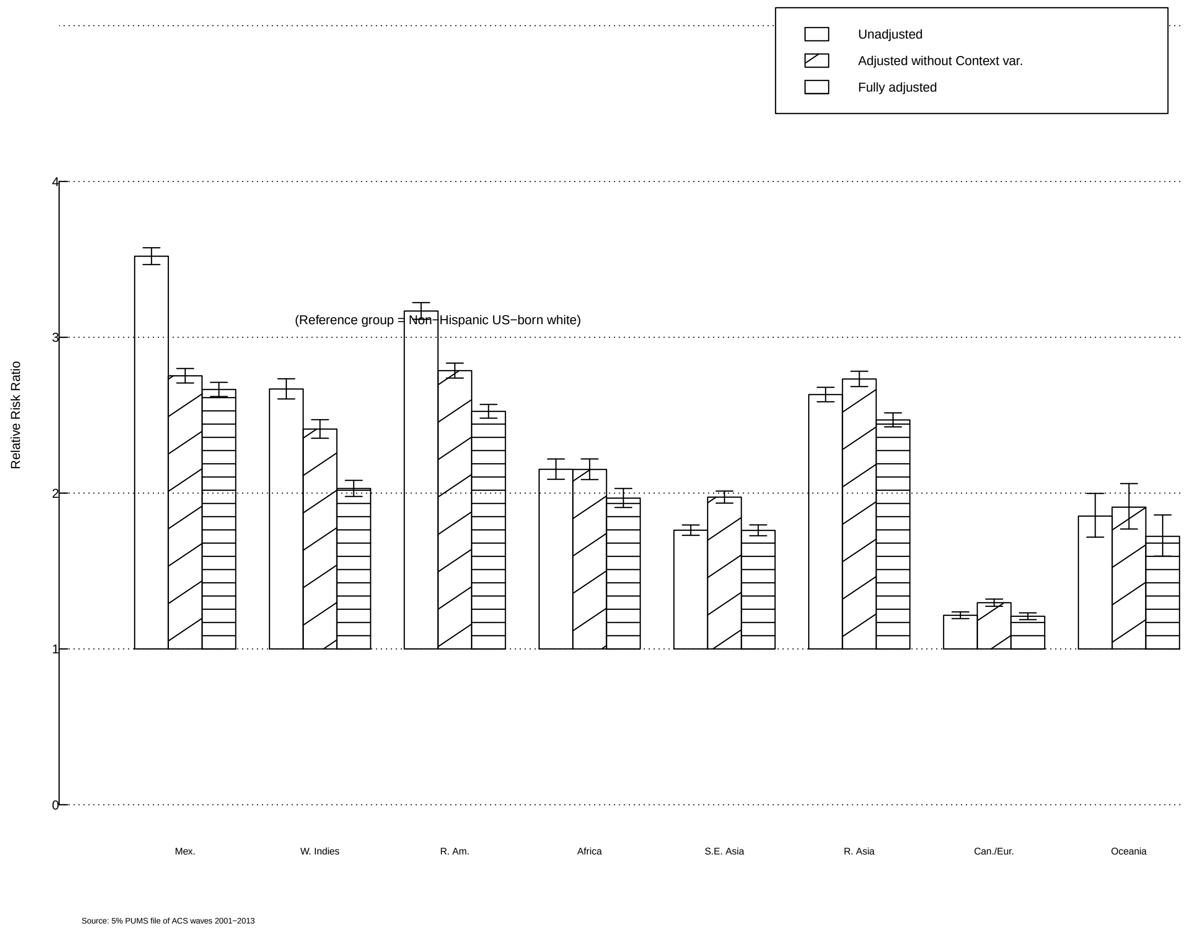


Figure 3.2: Relative Risk Ratios of Vertical Extension by Sending Region

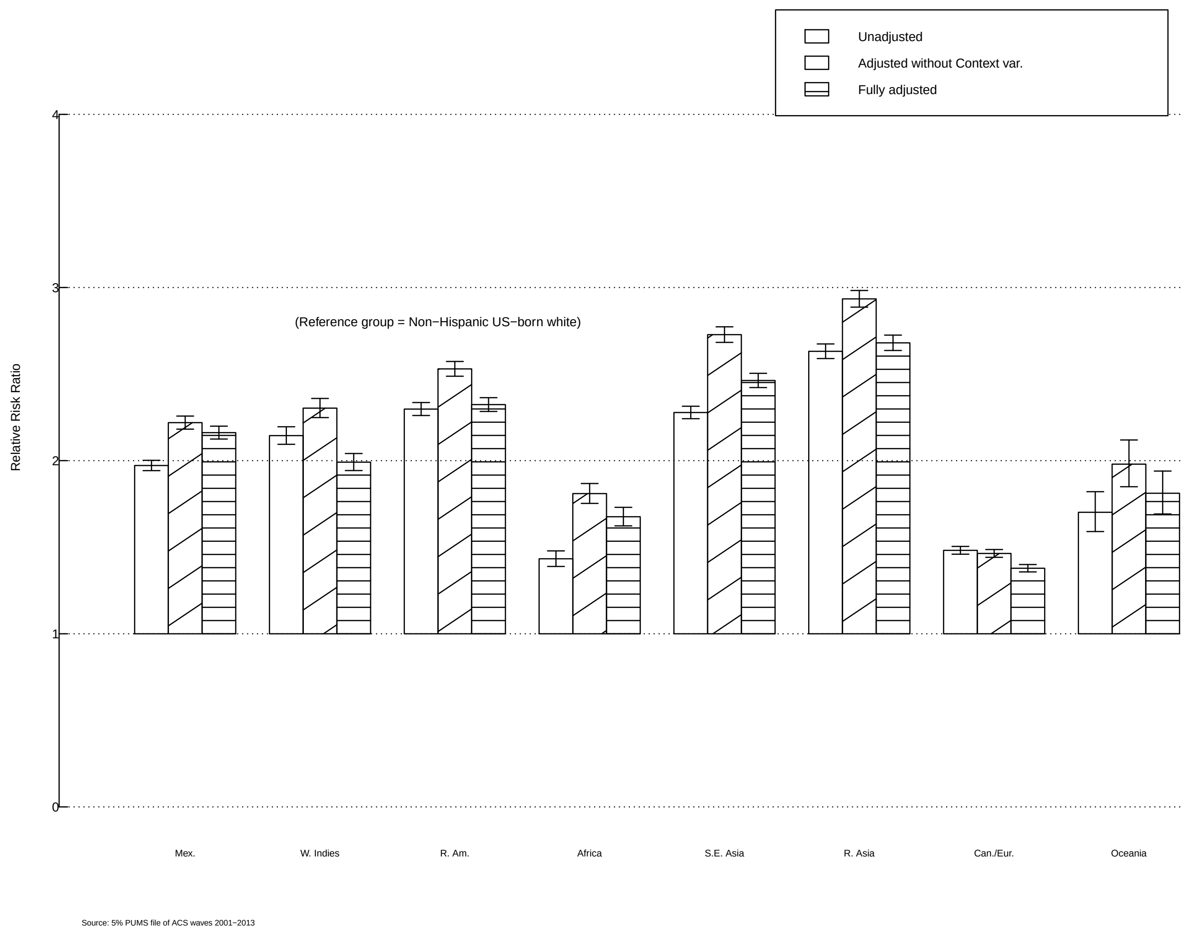


Figure 3.3: Fully Adjusted Relative Risk Ratios of Horizontal Extension by Income and Sending Region

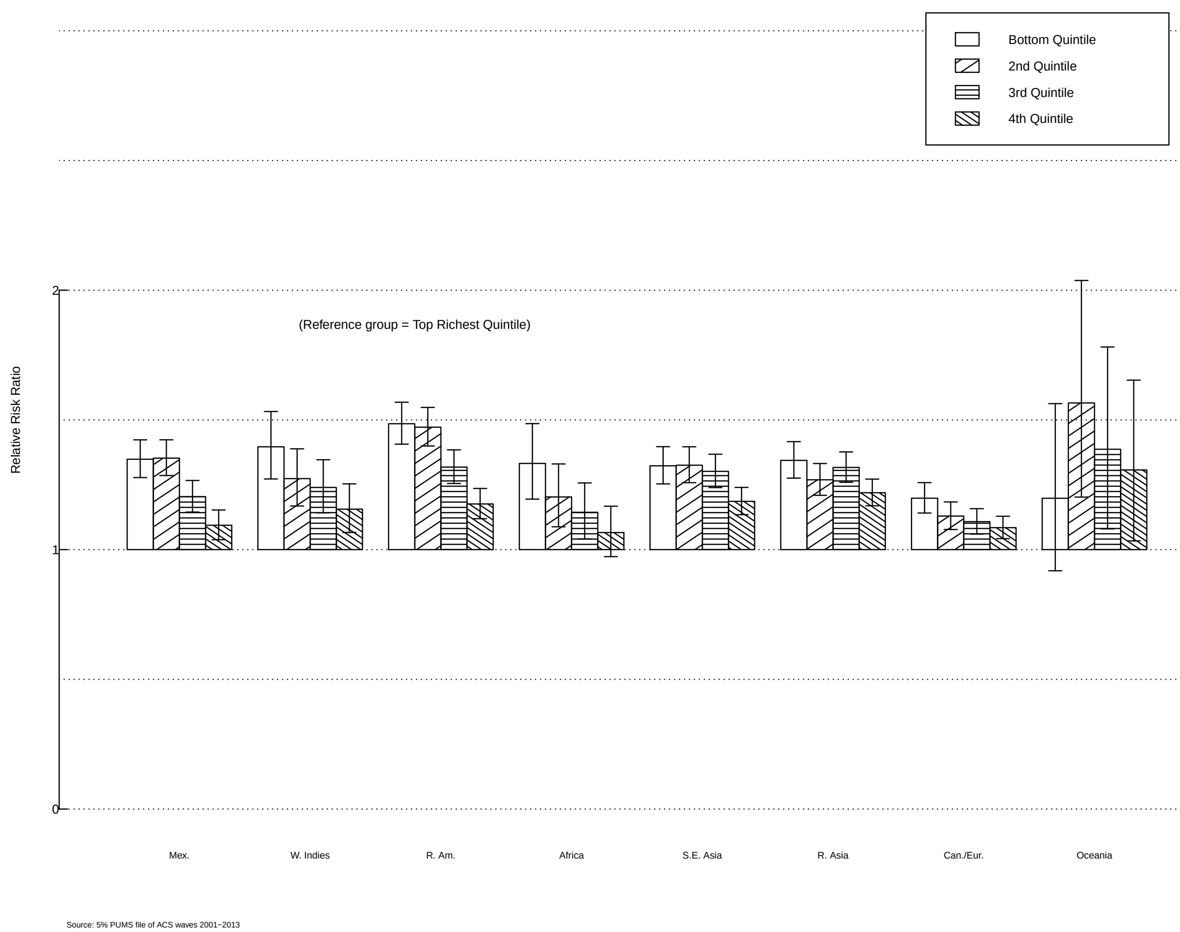


Figure 3.4: Fully Adjusted Relative Risk Ratios of Vertical Extension by Income and Sending Region

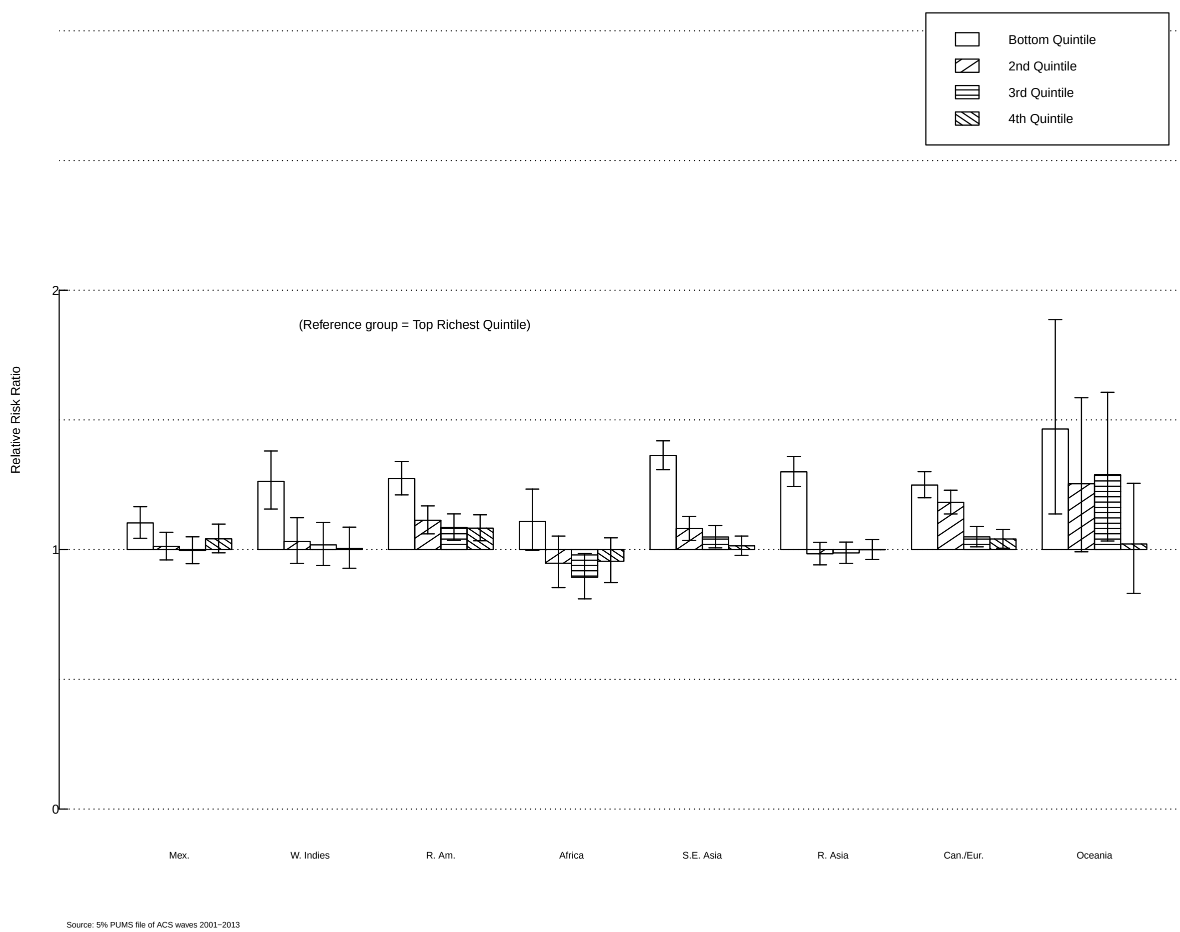
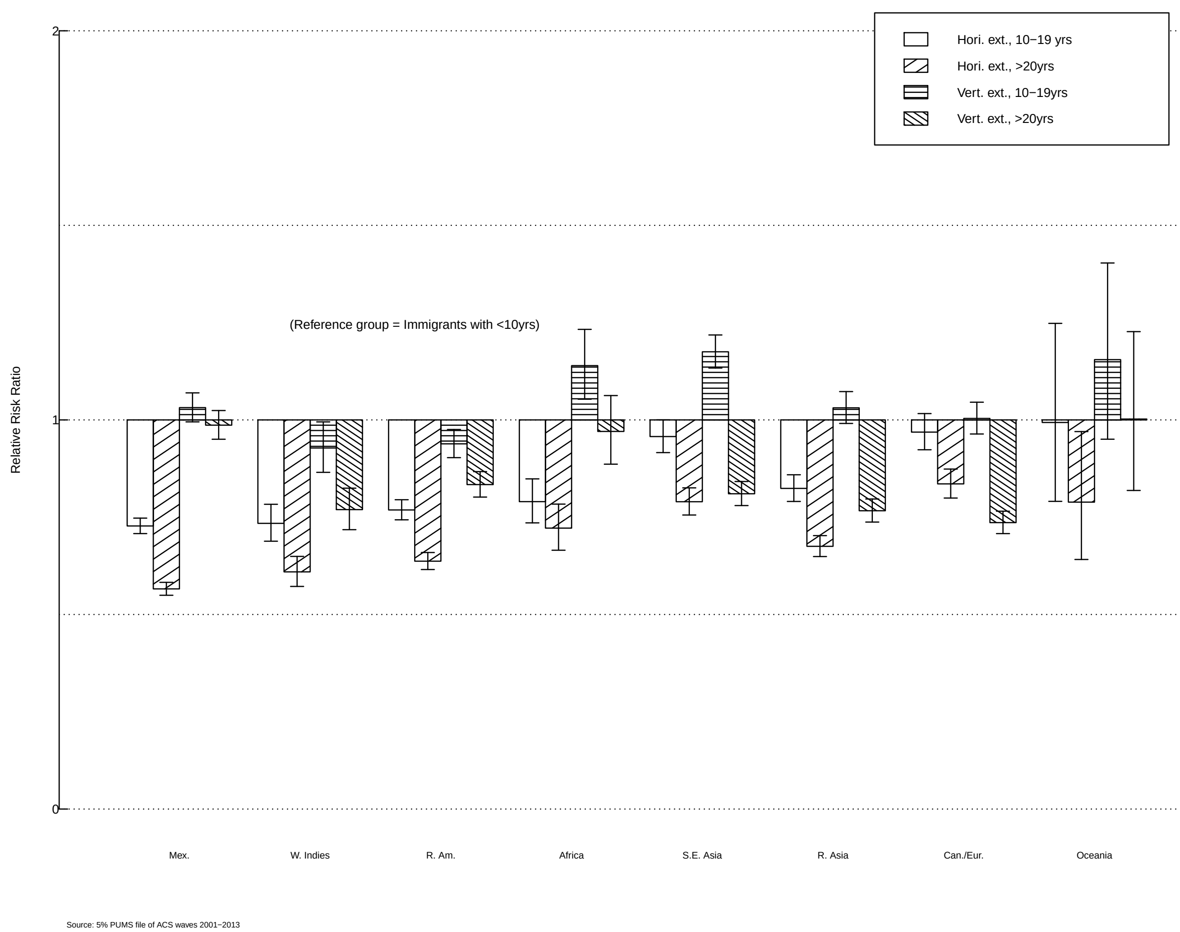


Figure 3.5: Fully Adjusted Relative Risk Ratios of Extension by Duration and Sending Region



CHAPTER 4 : Emigrant Selectivity and Choice of Migration Destination: Evidence from Ghana

Abstract

African emigration has been growing over the past few decades, but at the same time there is still little research on the nature and scope of selectivity among African emigrants to various destinations. In addition, because of a lack of data in most African countries, research on migration has not yet adequately investigated possible mechanisms whereby socio-demographic characteristics combine to shape migration decisions. Using data from the 2009 Ghana survey of the Migration between Africa and Europe (MAFE) project and data from nationally representative surveys in the United Kingdom and the United States, I investigate these research gaps among emigrants from Ghana. I find evidence for positive selection among current emigrants compared to both non-migrants and return migrants. My results further show that return migrants tend to have more favorable socioeconomic characteristics than non-migrants. My findings also point to selectivity among emigrants by destination such that emigrants who migrate to a destination further away are more positively selected. However, the results fail to show any effect of wage differentials at destination in explaining the choice of migration destination among emigrants to the United Kingdom and the United States. This paper offers three contributions to the literature on migrant selectivity. First, it focuses on selectivity among return migrants versus current emigrants; and emigrants across des-

tinations. Second, it assesses whether socioeconomic characteristics used at either the individual or household level influence migrant selectivity. Finally, it tests empirically two theoretical pathways relating migrant characteristics to the choice of migration destination.

4.1. Introduction

The rapid and sustained increase in emigration from Africa has heightened interest in the selectivity of these flows (Castagnone et al. 2015, Schoumaker, Flahaux, and Beauchemin 2015, van Dalen, Groenewold, and Schoorl 2003). Africa currently faces demographic and economic forces that are more powerful than those that led to mass emigration from pre-industrial Europe (Hatton and Williamson 2001, 2002). For example, most African countries experience rapid population growth and sluggish economic development, heightening the pressure to emigrate in search of opportunity (Hatton and Williamson 2001).

Understanding the nature and scope of selective emigration out of Africa is important for both theoretical and policy reasons, since it is closely related to the ongoing debates about brain-drain and its consequences for Africa (Anarfi, Quartey, and Agyei 2010, Docquier and Rapoport 2007, 2012). The outflow of skilled emigrants is not only thought to deprive African nations from citizens who can make the most valuable contribution to the development process, but it is also thought to represent losses in public spending on education as emigrants receive education in their origin countries before migration (Anarfi, Quartey , and Agyei 2010, Docquier and Rapoport 2007, 2012).

Although selective emigration out of Africa has important implications for development, it remains relatively understudied. As yet, because of data availability, there are four major knowledge gaps in our understanding of selectivity among African emigrants. First, most studies on the topic compare emigrants to the resident population of Africa, failing to account for the heterogeneity of the latter group, which includes both those who have never moved and those who previously emigrated but subsequently returned. Second, this literature provides few insights into how selectivity varies across migrant destinations. Third, research on selectivity often relies on either emigrant characteristics or characteristics of their household heads interchangeably to assess selectivity. Although one might expect some correlation between these measures, they potentially capture fundamentally different realities. Finally, the current literature does not often explore how migrant selectivity intersects with labor market outcomes in destination countries (Massey et al. 1993).

Using data from the 2009 Ghana survey of the Migration from Africa to Europe project (MAFE), the British Labor Force Survey (waves 2010-12), and the American Community Survey (waves 2005-10), this study bridges these gaps in our understanding of selectivity among emigrants from Africa. I focus on Ghana for three reasons. First, there are high quality data available on emigration from Ghana and on Ghanaian immigrants in developed countries. Second, Ghana experiences high rates of emigration, enhancing the salience of emigrant selectivity for development (Capps, McCabe, and Fix 2012). Finally, because of Ghana's cultural and economic proximity to most other Sub-Saharan countries, a study on emigration from this country will provide insights into emigration from the broader Sub-Saharan region.

This study first investigates how current emigrants compare to return migrants and non-migrants in terms of their socioeconomic characteristics. Second, it explores whether emigrants to various destinations differ in their socioeconomic characteristics. Third, this study examines selectivity by using alternative measures of socioeconomic characteristics at either the individual or household level. Finally, it investigates whether Ghanaian emigrants to high income countries base their migration decision on a comparison of earnings across destinations or whether they rely on the presence of networks to choose their emigration destination. This study develops models that control for several relevant variables used in the migration research. The paper offers three contributions to the literature on migrant selectivity. First, it focuses on selectivity among policy relevant groups (return migrants versus current emigrants; and emigrants across destinations). Second, it assesses whether socioeconomic characteristics used at either the individual or household level affect conclusions about selectivity. Finally, it tests empirically two theoretical pathways relating migrant characteristics to the choice of migration destination.

4.2. Background

4.2.1. Migrant Selectivity

Migration is often seen as a means to improve earnings opportunities of the working age population, lending itself to economic rationalization. Not surprisingly, neoclassical economic theories conceptualize migration as an investment in human capital (Harris and Todaro 1970, Sjaastad 1962, Todaro 1980), whereby a prospective migrant decides where to live by choosing among possible destinations, the one that provides him/her with the highest discounted earnings over the life course (Borjas 1994, Massey et al. 1993). In this formulation, earnings are a function of migrant human capital characteristics, such as educational attainment, and life course characteristics. Because their theoretical relevance strongly depends on whether various factor markets are complete, neoclassical economic theories do not fully explain forces that shape migration in the developing countries, where migration is thought also to operate as a risk management strategy that helps households diversify their sources of income (Senne, Chort, and Gubert 2011, Azam and Gubert 2006). In other words, migration serves to smooth household consumption and to improve a household's ability to invest through remittances (Castles, De Haas, and Miller 2014, Massey 1999, Massey et al. 1993, Stark and Taylor 1991).

Empirically, many studies have found a strong association between migrant human capital characteristics and the decision to migrate (Ackah and Medvedev 2012, Anarfi, Quartey , and Agyei 2010, Arthur 2012, Baizan 2014, Baizan and Gonzalez-Ferrer 2010, Black et al. 2013, Borjas 1994, Lucas 2013, Massey 1999, Massey et al. 1993). Specifically, it is well-documented that emigrants are typically in their prime working ages (Black et al. 2013, Schoumaker, Flahaux, and Beauchemin 2015, Funkhouser 2009, Hatton and Williamson 2001, Lucas 2013, van Dalen, Groenewold, and Schoorl 2003). These findings lend a strong support to the idea that migrants represent a selected subpopulation from their origin countries and do not represent the average person at either origin or destination. Comparing emigrants to their counterparts in their origin country, migrants tend to be better educated and wealthier than those who have never migrated or who have returned (Feliciano 2005, Jasso et al. 2004, Lucas 2013, Schoumaker, Flahaux, and Beauchemin 2015, van Dalen, Groenewold, and Schoorl 2003). The fact that immigrants must incur high migration costs together with the fact that they must comply with increasingly stringent immigration policies in receiving countries is thought to explain their positive selection based on socioeconomic characteristics (Czaika and Haas 2016).

Although several studies have found evidence of positive selectivity among emigrants, emigrant selectivity need not always be positive. As Borjas and Bratsberg (1996) have shown, the wage distribution along with the skill distribution in both the origin and destination countries shapes the type of selectivity observed in migrants. For example, when wages are highly unequal, less skilled (negatively selected) workers might find it attractive to emigrate (Borjas and Bratsberg 1994). Thus, for a given population, assessing the nature of selectivity is an empirical matter.

Another important finding in research dealing with selectivity among migrants is that it may evolve over time. As several studies have shown, given conditions in origin and destination countries, selectivity tends to vary by period (year) and cohort of migrants (Borjas 1994, Borjas and Bratsberg 1994). Changes over time in migrant selectivity can be explained by the network theory of migration, which holds that, for a given origin country, selectivity will diminish in subsequent migration flows because migrant networks (both at origin and destination) reduce the costs associated with migration (Castles, De Haas, and Miller 2014, Massey 1999, Moretto and Vergalli 2008, Taylor et al. 1989). Today, the network theory holds a prominent place in research on international migration. Family networks represent an important feature in current migration flows owing to the fact that most receiving countries support family reunification through legal means in their immigration policy.

The fact that migrants from a given origin move to multiple destinations has motivated further investigation of selectivity by comparing migrants across destinations. Scores of studies have found that the choice of the destination is related to migrant demographic and socioeconomic characteristics (Akee 2010, Fafchamps and Shilpi 2013, Funkhouser 2009). This suggests that the choice of the migration destination may be a manifestation of selection in a way that is consistent with dominant migration theories. For example, in a study of Ghanaian emigration to Europe, Castagone and colleagues (2015) find that the highly educated tend to migrate to the United Kingdom, which is Ghana's former colonial power, whereas the less educated often migrate to other European destinations (Castagnone et al. 2015). These authors further found evidence of heterogeneity across destinations among these migrants with respect to gender and family wealth (Castagnone et al. 2015). Their findings are consistent with the world systems (or globalization) theory of migration (Castles, De Haas, and Miller 2014, Massey 1999). This theory, drawing from Marxist views of the relationship between rich and poor nations and from dependency theory, postulates that colonial links explain a great deal of migration flows by creating proximity in terms of geography, social norms, and language, all of which lower migration costs (Castles, De Haas, and Miller 2014). Central to the world systems theory is the idea that migration is tightly connected to broader socioeconomic processes. As such migration is seen as a consequence of the disruption in social, economic, political, and cultural institutions caused by capital expansion in less developed nations (Castles, De Haas, and Miller 2014, Massey 1999, Massey et al. 1993).

Another preoccupation among researchers on migration has been to uncover mechanisms explaining observed patterns of selectivity among migrants. To address this problem, several studies have sought to investigate how pre-migration differences in human capital among migrants are related to their subsequent performance in the labor market in destination countries (Akee 2010, Fafchamps and Shilpi 2013, Funkhouser 2009). In this line of thought, Fefchamps and Shilpi (2011) investigate whether Nepalese migrants select the destination that results in the highest expected income by using a structural approach that relies on the neoclassical migration theory. Contrary to most theoretical predictions, they find that differences in expected income holds little sway in explaining migrants' choice of destination, while ethnic similarity, a measure of migration costs or networks, plays a more important role.

4.2.2. Return migration and selectivity

Despite the fact that the decision to migrate is often modeled as a one-shot decision, there is an increasing awareness among researchers that migration is not always a permanent move, resulting in an increased interest in return migration. The study of return migration is also motivated by the fact that long-run comparative analyses of migrants with non-migrants requires a better understanding of how immigrant stocks at different points in time compare to initial entry flows, because selective return migration can distort any comparison over time (Constant and Massey 2004, Flahaux 2015, Flahaux 2016). Conceptually, there are two possible paths to return migration. First, return migration may be the result of ex-ante calculations. Some migrants may have specific and predetermined goals in terms of savings or skill accumulation (Borjas 1994, Massey et al. 1993). After achieving these goals such migrants return home. On the other hand, ex-post adverse shocks at destination may push migrants to repatriate (Flahaux 2015). These migrants fail to achieve economic success because of either low achievements in the job market in the destination country or higher migration costs than initially expected (Duleep 1994).

An alternative theory of return migration draws upon a rational choice framework (Dustmann, Fadlon, and Weiss 2011, Dustmann and Glitz 2011, Dustmann and Kirchkamp 2002, Dustmann and Weiss 2007). From this perspective, return migration might first come about if migrants value the place where their consumption takes place. Such a spatial preference in consumption encourages return migration even if wage differentials persist (Dustmann and Weiss 2007). Second, differences in purchasing power between the origin and destination countries can possibly help migrants accumulate capital (skills or financial assets) at destination, which can prove helpful in enhancing their living conditions after their return to their origin country. Finally, when rapid industrialization is taking place in the origin country, migrants may gain from transferring their newly acquired skills back home, where they would reap higher returns than in the destination country (Dustmann and Weiss 2007).

By its very nature, return migration can lead to selection. As with migration, it is hard to predict the nature of selectivity that prevails among return migrants, since this selectivity depends entirely on the mechanisms shaping the process of return migration. Several studies on return migration, conducted in destination countries, have found evidence of various forms of selectivity. While Jasso and Rosenzweig (1988), with cross-sectional data, find that return migrants are positively selected (Jasso and Rosenzweig 1988); other researchers using longitudinal data find negative selection among return migrants (Borjas 1989, Lindstrom and Massey 1994, Massey 1987). Yet another group of studies found no evidence of selectivity whatsoever among return migrants (Chiswick 1986, Reagan and Olsen 2000). With studies using data from the origin countries, results are also mixed. Exploiting the absence of selectivity among Albanese immigrants in Germany (Constant and Zimmermann 2003), Gaule (2000) finds evidence of negative selectivity among return migrants relative to those who have never migrated (Gaul 2014). Similarly, Wahba (2015) finds evidence of negative selectivity among return migrants in Egypt (Wahba 2015). However, Ambrosini and colleagues (2015) document positive selectivity among return migrants in Romania (Ambrosini et al. 2015). The findings suggest that the nature of selectivity among return migrants varies across contexts.

4.2.3. New Directions in Research on Migrant Selectivity

Although past studies have enhanced our understanding of selectivity among emigrants, there still remain four gaps in the current literature on emigrant selectivity. First, past research does not often distinguish between non-migrants and return migrants when migrants who have not returned are compared to populations in sending countries. Most studies on migrant selectivity compare emigrants to the resident population in the origin country without accounting for the fact that this population includes both people who have never moved and those who have come back after time abroad. Because return migrants are likely to differ from non-migrants in ways that are difficult to anticipate, results from the analysis of selectivity may be biased. Current estimates of selectivity are likely to underestimate emigrant selectivity as return migrants may be more similar to current emigrants than to those who never migrated and have higher human capital than non-migrants irrespective of the selection mechanisms driving return migration.

Second, most studies of migrant selectivity do not compare across destinations. Given that financial and legal constraints on immigration vary across destination countries, one would expect various migration flows to be differentially impacted by such constraints. In other words, to the extent that emigration to neighboring countries is less financially and politically constraining than emigration to distant countries, it can be postulated that the former should be less selective than the latter. In addition, because the need for skills in the destination country varies according to its stage in economic development, migration flows to different destination countries are likely to be heterogeneous.

Third, past research often fails to investigate mechanisms driving observed patterns of selectivity among migrants. How differences in migrant profiles are related to expectations about economic performance in the destination country is not yet clearly understood. A microanalysis involving an approach that seeks to test mechanisms explaining selectivity is needed to improve our understanding of the processes of migration (Massey et al. 1993).

Finally, past studies have used emigrants' socioeconomic characteristics and those of their household heads interchangeably. Although both sets of measures are likely to be correlated, they may well tap different dimensions of socioeconomic status. While the former captures emigrants' characteristics related to human capital, the latter provides information about socioeconomic background. Results from studies using one set of measures might differ from those using the other set of measures, providing a possible explanation of the discrepancies across studies.

This paper addresses the limitations of previous studies in four principal ways.

First, I use the MAFE data to compare current emigrants with return migrants and non-migrants. My focus is twofold, as I seek to investigate how current emigrants compare to non-migrants and return migrants, and how return migrants and non-migrants compared to one another. Second, grouping emigrants by destinations, I am able to compare the profile of emigrants across destinations and examine whether there is evidence of differential selectivity. Third, combining data from the American Community Survey, the Labor Force Survey in the United Kingdom, and the Migration from Africa to Europe, I predict for each emigrant, the expected earnings he/she could expect at each destination based on the emigrants' observed characteristics and then investigate how the actual choice of destination relates to the expected earnings at destination. At the same time, I explore the relative merit of the neoclassical economic theory and the network theory of migration. Finally, I use both emigrant's education and education of the household head to investigate selectivity.

4.3. Data and Methods

The data are mainly drawn from the Ghana survey of the Migration from Africa to Europe project, fielded in 2009. They were obtained from the Institut National d'Etudes Demographiques (INED) in France. The MAFE project was designed to collect data to gain a better understanding of migration between Africa and Europe. The survey provides detailed socioeconomic information about non-migrants, return migrants and emigrants. In the MAFE dataset, emigrants are defined as household members who have moved abroad and who either intend to live permanently there or who have lived continuously for at least six months outside of Ghana; non-migrants refer to individuals who have never lived outside of Ghana; and return migrants have returned to Ghana after living abroad for at least six months. Information about most emigrants, non-migrants, and return migrants is provided by proxy respondents. The design of the survey is similar to the Mexican Migration Project.[1] The MAFE survey collected sociodemographic information, such as educational attainment, relationship to household head, and age, at the individual and household levels in Ghana. The survey also documented migration histories, including migration duration for each migratory episode for most adult family members (who were living in Ghana or in Europe) (Beauchemin et al. 2014).

The MAFE project is a relatively small scale survey and is not nationally representative of the Ghanaian population (Beauchemin et al. 2014). However, it is representative of the Kumasi and Greater Accra regions (Beauchemin et al. 2014), which are the two most populous among the ten regions in Ghana and the main sending regions of emigrants to Europe and the Americas (Ghana Statistical Service 2012). I restrict the sample to household members in their prime working ages (25-64 years old). The restriction of the sample to this age group is motivated by the fact that I am interested in education selectivity into emigration, which requires that the household members have completed their educational attainment. I further restrict the dataset to cases with no missing information on any explanatory variable. Cases with missing information represent 3.7, 0.5, and 4.8 percent of the non-migrants, return migrants and emigrants, respectively. This leads to discarding 3.7 percent of the initial sample. The final analysis file consists of 3,837 Ghanaians in their prime working ages, including 1,019 emigrants in various destinations[2] and 382 return migrants.

Because information on earnings abroad is not available in the MAFE dataset, I draw on data from two major destination countries, the United Kingdom and the

[1]http://mmp.opr.princeton.edu/research/studydesign-en.aspx

[2]This relative share of emigrants is high. In all probability, this may stem from the fact that relatives who move abroad are considered household members while those who have moved out the household but who still live in Ghana are not accounted for in the listing of household members.

United States, to test whether earnings differentials shape the choice of the migration destination. First, I use five waves (2005-10) of data from the American Community Survey conducted by the US Census Bureau. The ACS data provide information about annual earnings in the United States, along with their human capital characteristics. I restrict the ACS sample to people whose country of birth is Ghana and who are in their prime working ages (25-64). I further restrict this sample to immigrants who are currently working either as an employee or self-employed because earnings are only defined for those who are economically active. The pooled ACS data yield an analysis sample of 5,823 Ghanaian immigrants in the United States. For earnings, there are missing cases, representing 48.0 percent of the sample. I impute this missing information on migrant's earnings by using a multivariate normal imputation procedure with the number of hours worked as an ancillary variable.

Second, the labor market data of Ghanaians in the United Kingdom are drawn from twelve quarterly waves[3] of the British Labor Force Survey (LFS), fielded by the British National Office of Statistics. The LFS is the largest UK household survey, and it is based on a nationally representative sample. It provides employment data and rich socioeconomic information about people. I restrict the sample to the 809 Ghanaian immigrants who are in their prime working ages (25-64) and working (as an employee or a self-employed) at interview. In this dataset, the earnings variable has missing cases, representing 79.4 percent of the sample. A multiple imputation procedure with a multivariate normal variable covariance structure helps handle such cases with missing information (See Tables A4.1 and A4.2 in the Appendix).[4]

[3]January-March quarters of 2010- October-December quarters of 2012.

[4]No ancillary variable is used for this imputation, because the number of hours worked also has missing information.

4.3.1. Dependent variables

Four measures, referring to migration status, are used as dependent variables. First, a binary variable indicating whether an individual has been living abroad for at least six months or whether an individual, who has moved, intends on staying abroad permanently is used to capture the migration status of the housheold members. Second, because this indicator does not help distinguish between return migrants and non-migrants, I construct a three-item categorical variable, indicating whether a household member is a current emigrant (who intends to stay abroad), return migrant, or non-migrant. Third, a four-item categorical variable is created to indicate the four possible destination regions among emigrants. Such regions consist of Africa, Europe, the Americas, and Other Destinations (Asia and Oceania). The definition of these regions of destination is based on the historical context of migration outside Ghana, which is marked by geographical, cultural, and institutional ties to different destinations. Unfortunately, the small sample size precludes my ability to distinguish between the countries that are member states of the Economic Community of West African States and the other African countries. This third dependent variable helps investigate emigrant selectivity across destinations. Finally, a binary indicator is used to distinguish the migration destination among emigrants to the United Kingdom and the Unted States.

4.3.2. Explanatory variables

I use several explanatory variables based on neoclassical economic theory, network theory, and life course perspectives on migration. People's education and the education of the household head are the key socioeconomic explanatory variables. They are defined as categorical variables, which are constructed from a variable that captures the highest education attained by each individual. I define five indicators distinguishing between no education, primary education, middle or junior high school, secondary or senior high school, and higher education. Because the MAFE data do not provide information about earnings in the United Kingdom and the United States, I predict values of expected earnings with data from the ACS (United States) and the LFS (United Kingdom) restricted to Ghanaian nationals in each destination country. Hence, expected earnings are used as predicted explanatory variables. It is important to highlight that owing to the fact that earnings are expressed in current local currency, proper adjustments are made to account for differences in currencies and living standards between the United States and the United Kingdom. In a first step, such adjustments consist of using the purchasing power parity converter developed by the World Bank's International Comparison Program to convert UK earnings into current US dollar. I next use the consumer price index, indexed at years 1982-84, to account for inflation in migrant earnings.

Because migration decisions often require a substantial investment, relationship to the household head has often played an important role in affecting whether people migrate. For this reason, I use five binary variables to control for the relationship to the household head (household head, spouse of household head, child of household head, sibling of household head, and other relationship to household head). Another set of explanatory variables consists of life course variables, which have proven important in explaining migration decisions. Such variables include three binary indicators for the household member's age group at the time of the survey (25-34; 35-49; and 50 and above); an indicator of whether the household member is male; and a variable indicating whether the household member is married at the time of the survey. Although most of these variables are not captured at the time of migration (except age), it is unlikely that the results would be seriously biased because most people in the dataset migrate in their thirties. I further control for household size and for networks in Ghana and abroad by including a variable indicating whether the household member lives in Accra versus Kumasi and the number of household members who are currently living in Europe, Africa, and other destinations. Finally, among emigrants, I control for the period of emigration, by using three binary indicators (before 1990, 1990-2000, and 2000 onward).

4.3.3. Empirical Strategy

Following previous studies on migration, I begin by modelling the migration decision as a function of observable characteristics of non-migrants, migrants, and return migrants. A series of logistic regressions are estimated to predict the odds of migration. In a subsequent step, I distinguish between non-migrants and return migrants. Because the outcome variable is a three-item categorical variable with no a priori ordering, I estimated a series of multinomial logit regression to predict the odds of each migratory status. To explore the structural relationship between migrant human capital and their migration destination, I further estimated a series of logistic regressions predicting the odds of choosing the United States as a destination relative to the United Kingdom. It is important to note that I account for the unequal sampling probability of each household by weighting all estimates by appropriate weights provided by the data providers (INED, UK's National Office of Statistics, and US Census Bureau). Finally, I compute robust standard errors to address issues related to clustering at the household level. All estimates are expressed as odds ratios.

In Model 1, I control only for household member's education and age. This base model seeks to examine the unadjusted contribution of education to the migration outcome, whether it is the emigration decision or the decision related to the choice of the migration destination. Model 2 adds the education of the household head to Model 1 to help dissociate the effect of family socioeconomic status from those directly related to emigrant's own educational attainment. On the assumption that there is social reproduction, most studies have used either measure to capture household member's socioeconomic status. But it might be the case that these two measures are tapping different dimensions of socioeconomic status, and hence may be differently affecting the migration outcomes. Model 3 adds the set of socio-demographic determinants of migration that have been used in most empirical analysis of migration. These include gender, marital status, relationship to the household head, number of household members living abroad in various destinations, and region of origin in Ghana.

To further investigate whether there is selectivity into different destinations, I restrict the analysis to emigrants classified by their regions of destination. This specification allows for the use of the year of emigration as an additional variable in the model. However, because the sample size is very small, I am not able to explore simultaneously the contribution of both the educational attainment of the emigrants and that of their household head. Hence, I use a modified version of Model 3 that includes only the educational level of the emigrants. Model 4a adds the full set of other socio-demographic characteristics to Model 1, while Model 4b adds this set of covariates to a model that includes the educational attainment of the household head rather than the individual's own educational attainment.

Finally, because I am interested in investigating whether wages at destination influences the choice of migration destination of Ghanaian emigrants to high income countries, I focus on Ghanaian emigrants to the United States and the United Kingdom only. Based on the MAFE data, Europe and the Americas account for approximately 87 percent of the Ghanaian emigration stream (author's own calculation based on the MAFE). Although each European country could be a potential destination, past research on emigration from Ghana to Europe has shown that three European countries (the United Kingdom, Germany, and the Netherlands) receive most Ghanaian emigrants (Schmelz 2009). Similar results are found from the MAFE data, indicating that the United Kingdom accounts for more than 52 percent of the Ghanaian emigrants to Europe, with the next European country (Italy) accounting for at most for 15 percent. When it comes to the Americas, the United States accounts for more than 85 percent of Ghanaian emigration to the Americas (author's own calculation based on the MAFE).

I then predict the likelihood of emigration to the United States as a function of the difference in returns to emigration between the United States and the United Kingdom, where returns to emigration are captured by the expected earnings of emigrants (Fafchamps and Shilpi 2013, Greenwood 2005). These expected earnings are modeled as a function of emigrant characteristics such as schooling,and age (at emigration).

Because the MAFE data do not provide information about earnings in the United Kingdom and the United States, I use a two-step approach to generating the mean expected earnings that migrants would attain in each destination (Fafchamps and Shilpi 2013). First, I estimate the parameters of earnings regressions with data from the ACS (United States) and the LFS (United Kingdom) restricted to Ghanaian nationals in each destination country. Specifically, using the human capital approach, I estimate a regression of earnings of Ghanaian immigrants based on their educational attainment, age at emigration, age at emigration squared, and gender (Borjas 1994, Chiswick 1978, Massey et al. 1993).

$$y_j^i = \alpha_i + \delta_i X_j + \epsilon_j^i \tag{4.1}$$

Where:

- i = 0 (UK) or i = 1 (US).
- X_j are observable human capital characteristics of migrants

Note that the dependent variable is log-earnings and δ_i is freely estimated to reflect the fact that returns to observables may vary by destination.

Equation (1) specifies the earnings regression, which is estimated separately for each country, using the sample of employed Ghanaian immigrants. Second, the estimated parameters (coefficients) in Equation (1) for each explanatory variable are recovered and used to predict the expected earnings of each emigrant in either the United States or United Kingdom based on the same individual characteristics in the MAFE dataset. It is important to highlight that the coding of human capital variables (education and age at emigration) is kept the same across countries. There are potential problems with using contemporaneous human capital characteristics of emigrants to predict their expected earnings (except age), as these characteristics could have changed since the time they emigrated. Unfortunately, the MAFE data do not specify when and where education is completed (I shall defend the validity of using the education variable in the result section) (Appendix Tables A4.3 for the earnings regression results).

The model specified in Equation (2) is the central piece of my empirical strategy. It regresses migration destination on differences in expected earnings and other explanatory variables available in the MAFE data. Such explanatory variables include the year of emigration, used to control for the macroeconomic characteristics such as business cycles. To account for heterogeneity stemming from the region of origin in Ghana and for migrant domestic networks, I use an indicator of the sending region. Migrant networks in the destination country are accounted for by using the number of household members who are currently living in Europe, in Africa, and elsewhere (Americas, and Asia/Oceania)[5].

$$Pr(i=1) = F\big(\rho E(y_j^i - y_j^0) + \gamma Z\big) \tag{4.2}$$

$$= F\big(\rho\big(\alpha_1 - \alpha_0 + (\delta_1 - \delta_0)X_j + (\beta_1 - \beta_0)U_j\big) + \gamma Z_j\big) \tag{4.3}$$

Where:

- F(.) is chosen to be a logistic link function and Z is the set of control variables for migrant j;
- ρ is the odds of the difference in expected earnings in the United States and in the United Kingdom;
- γ and is the vector of the odds of associated with each control variable (region, year of emigration, gender, marital status, household size, number of household members in various destination).

The approach, described above, seeks to examine a structural relationship between the choice of the migration destination and migrant characteristics that operate through their returns to skills in the destination country. An important implication of this approach is that the contribution of emigrant characteristics to the decision making process is only captured through their effects on expected earnings at destination.

[5]While these numbers include current emigrants at different destinations, they do not include household members in Ghana.

With this approach, I allow the relationship between emigrant skills and their expected earnings to differ by receiving country to capture differences in the way skills are rewarded in each country. But this specification maintains the assumption that the skills have the same ordering in both receiving countries. Because migration costs are perfectly collinear with the migration destination, I cannot account separately for these costs. In other words, net migration costs are absorbed in the intercept of the wage regressions. As such, the estimates of returns should be interpreted as gross returns to migration.

4.4. Results

Table 4.1 summarizes socioeconomic and demographic characteristics by whether housheold members are current emigrants, return migrants or non-migrants . Return migrants represent 10 percent of the sample, which is a relatively high rate of return migration. To put this number in context, a recent study of return migration in Romania defines shares of return migrants ranging between 7 and 9 percent as high (Ambrosini et al. 2015, Flahaux 2015). Based on these figures, one would expect that about 27.3 percent of current emigrants would return to Ghana at some point in time.[6]

4.4.1. Migration, Return Migration, and Socioeconomic Selectivity

Table 4.1 further highlights notable socioeconomic differences among individuals across the various migrant categories. Current emigrants and return migrants are more likely than non-migrants to have a household head with a higher education. While 24 and 27 percent of individuals, respectively, from these former categories have a

[6] $27.3 = 100\text{x}\frac{382}{1019+382}$

household head with higher education, only 17 percent of non-migrants live in households headed by someone who has attained higher education. A different pattern is observed in people's own education. While return migrants are less likely than non-migrants to have completed schooling between primary and secondary education, both current and return migrants tend to have attained higher education than non-migrants. About 35 and 31 percent of current and return migrants have completed higher education, respectively, compared to 18 percent among non-migrants. On the other hand, non-migrants are more likely to have completed primary, middle and secondary school education than both current and return emigrants. Surprisingly, there is a high percentage (19 percent) of return migrants with no education, compared to 6 and 4 percent among non-migrants and current emigrants respectively. These bivariate results suggest negative selectivity among return migrants compared to both non-migrants and current emigrants.

Also evident in the table is heterogeneity across migrant categories with respect to most socio-demographic characteristics. For instance, the mean age is lowest among non-migrants (38 years old) and highest among return migrants (43 years old). Current emigrants have their mean age between these two extremes (41 years). Age at the time of migration (30.2 among current emigrants) is clearly lower than the age at interview. This mean age suggests that most migrants would have completed their schooling and married before migration.[7] Employment rates are higher among current emigrants compared to non-migrants and return migrants, with the latter two categories experiencing similar rates. While return and current emigrants are more likely to be male, non-migrants are more likely to be female. The household size differs across categories. In terms of timing of emigration, most current migrants have

[7]In Ghana, the median age at first marriage in Ghana is 19.8 and 25.9 among women aged 25-49 and men aged 30-59, respectively (2008 Ghana Demographic and Health Survey), retrieved at: http://www.dhsprogram.com/pubs/pdf/SR172/SR172.pdf

moved after the year 2000. The results suggest an increasing rate of migration over time. Although this could stem from a real increase in the propensity for emigration, it might also stem from a change in how household membership is defined over time. In other word, the results might be consistent with the scenario that members who stay away for a long time period are less likely to be considered a household member even if there is no increase in the emigration rates over time.

Table 4.2 presents summary statistics of emigrants by emigration destination. In terms of migration destinations, 47 percent of Ghanaian emigrants reside in Europe, followed by 41 percent in the Americas, 10 percent in Africa, and 3 percent in Asia and Oceania. The United States is the single most common destination country with 35 percent of all Ghanaian emigrants. Furthermore, the United States accounts for 85 percent of emigrants to the Americas. In Europe, the United Kingdom is the primary destination with 52 percent of Ghanaian emigrants to Europe (Results noot shown).

One can see large socioeconomic differences among emigrants across various destinations. Emigrants to Europe and the Americas are the most positively selected with respect to the education of household head, while those to Africa are the least positively selected. Household heads of emigrants in the latter category are more likely to have no education. Emigrants to other destinations (Asia and Oceania) are more likely to come from households whose heads have intermediate educational attainment. As far as emigrants' own education is concerned, emigrants to Asia and Oceania are the most likely to have completed higher education compared to the other emigrants. However, the small sample size (30 observations) precludes drawing firm conclusions about this pattern. Among the three remaining destinations, emigrants to the Americas average higher levels of educational attainment than those to Europe and Africa (41 percent versus respectively 31 and 13 percent). Current employment is higher among emigrants to Europe and the Americas compared to those in Africa and Asia and Oceania. The proportion of males is the highest among emigrants to Europe, and the lowest among those to Asia and Oceania. There are some signs of a sharp increase in emigration from Ghana to the relatively more developed countries (Europe, the Americas, Asia and Oceania) after 1990.

Another interesting pattern in the data hints at evidence of family networks influencing the choice of migration destinations, since migrants from a given household tend to move to the same destination. For example, the mean number of migrants to Europe is 1.9 among households with emigrants to Europe; and for households with members in Africa, this mean number is 1.7. Likewise, the mean number of migrants to the rest of the world (consisting of the Americas, Asia, and Oceania) per household is 2.1 and 1.7 among households sending migrants to the Americas and Asia and Oceania, respectively. In addition, the choice of the destination country seems to be related to the region of origin in Ghana. The majority of emigrants to the Americas (56 percent) and Asia and Oceania (61 percent) originate from the Greater Accra region, while emigrants to Europe mostly come from the Kumasi region (52 percent). There are also age differences in emigrants across destinations. Emigrants to the Americas are slightly older (42 years old) than emigrants to Europe (40 years old), Africa (41 years old), and other Asia and Oceania (41 years old).

Following the standard approach in the migration literature, I first examine both unadjusted and adjusted differences in socioeconomic characteristics between emigrants and a category that combines non-migrants and return migrants. Table 4.3 presents odds ratios from logistic regression models of a measure indicating whether household members are currently living abroad. The first column of the table pertains to Model 1, which produces unadjusted odds ratios of household member's own education. Consistent with the bivariate analysis, higher levels of education of the individual are generally associated with higher odds of having emigrated, though all the coefficients are not statistically significant. As seen in Table 4.3, higher education increases the odds of emigration by a factor of 3.7 compared to those with no education. The odds of living abroad are the highest among individuals aged 35-49. The next column for Model 2 adds the education of the household head. This leads to some attenuation of the association with one's own higher education, though it remains statistically significant. These results show selectivity based on the education of both the emigrants and their household head. The last column, for Model 3, adds controls for demographic and life course characteristics. Men are more likely to be living abroad than women. Married household members are also more likely than their non-married counterparts to have moved out of Ghana. Being the child of the household head is associated with higher odds of emigration, although the coefficient is not statistically significant. At the same time, households are more likely to have relatives abroad (siblings and other relatives). And the larger the household, the higher the odds of having an emigrant household member. The presence of household members abroad is associated with higher odds of emigration. The odds of living abroad increase by a factor of 3.3, 2.6, and 2.4 with an additional household member in Africa, Europe, and the rest of world, respectively.

I next examine differences in socioeconomic characteristics between emigrants, non-migrants, and return migrants accounting for demographic and life cycle characteristics. Table 4.4 shows odds ratios from multinomial logistic regression models of non-migrant, return migrant, and current emigrant status. The first two columns of the table pertain to Model 1, which produces unadjusted odds ratios of household members' own education relative to the non-migrants controlling for age at the time of survey. The results suggest important differences in the selection mechanisms between return and current emigrants. While higher education is associated with higher odds of living abroad relative to non-migrants, the magnitude of the association is reversed for return migrants relative to non-migrants. These results highlight a positive selectivity based on education among current emigrants. They also indicate that return migrants tend to be have lower levels schooling than non-migrants, which could not be inferred from Table 4.3.

The next two columns present results for Model 2 that adds the education of the household head to Model 1. Two sets of results emerge for current and return emigrants. First, adding household head's education reduces but does not eliminate the effects of household member's own education on the odds of being either current or return migrants. Among return migrants, while there is little difference in coefficients for household members' own education relative to no education, there is striking evidence of positive selection with respect to the education of household heads relative to non-migrants. Among current emigrants, the association of household members' own education is only modestly reduced by the inclusion of the education of the household head. Yet, there is evidence of positive selectivity among current emigrants compared to non-migrants based on the education of the household head. Second, selectivity based on the education of the household head is more pronounced among return migrants than current emigrants, as the odds ratios are larger and have higher statistical significance (the difference in the odds is statistically significant).

The last column, for Model 3, controls for demographic and life course characteristics. The patterns seen in the two previous columns are roughly unchanged. The role of demographic and life course characteristics in shaping return migration is mostly similar in direction, but different in magnitude from that observed among current emigrants. The various indicators of the relationship to the household head display no association with the odds of return migration. However, return migration is relatively more common in the Greater Accra region than in the Kumasi region.

Finally, network effects, measured by the number of household members in various destinations, remain strong and positive even after controlling for other socioeconomic characteristics.

4.4.2. Selectivity Across Destinations

The pronounced differentials in selectivity across migration statuses invite further examination of the selection patterns by destination. Hence, I estimate the odds of migration to each migration destination relative to migration to Europe. To investigate this question, I drop the subset of emigrants to Asia and Oceania because their small number leads to unreliable (unreasonably large) odds ratios. Furthermore, because I am dealing with a relatively small sample, I use education of either the migrant or household head as a measure of socioeconomic status. The full model here adds years since emigration along with the set of demographic and life course characteristics of emigrants to Model 1 or its variant that uses education of the household head. Table 4.5 presents odds ratios for emigration to the Americas and Africa relative to Europe separately. As shown in the Table 4.5, based on emigrants' own education, there is evidence of positive selection among emigrants to the Americas relative to those to Europe although the coefficients are not statistically significant, whereas emigrants to Africa appear to be negatively selected compared to those who emigrate to Europe and to the Americas.[8] These emigrants tend to have lower educational attainment compared to emigrants to Europe and the Americas. The estimates indicate no selection based on age across destinations, with emigrants to Africa being slightly younger than emigrants to either Europe or the Americas. The last two columns of the table present results that control for the education of the household

[8]There is a statistically significant difference between emigrants to Africa and those to the Americas (results not shown).

head instead of emigrants' own education. As seen in the case of individual's own education, there are no differentials in the effect of the education of household head among emigrants to the Americas compared to those to Europe, while the emigrants to Africa are, once again, negatively selected compared to emigrants to Europe and the Americas. As shown above (in Table 4.2), emigrants to the Americas tend to move from relatively large households compared to those to Europe, whereas emigrants to Africa do not present any difference with emigrants to Europe with respect to household size. As expected, the number of household members in Europe reduces the odds of emigration to either the America or Africa, while an additional household member in Africa is associated with higher odds of emigration within Africa. Likewise, an additional household member in the rest of the world (the Americas, Asia, and Oceania) is associated with a higher odds of emigration to the Americas. This set of results strongly suggests that networks affect the choice of emigrant destination.

4.4.3. Destination Choice, Expected Earnings, and Networks

To investigate the possible selection mechanisms in the neoclassical economic framework, I next focus on emigrants to the United Kingdom and the United States. I begin with some background information about emigration streams into these two destinations. Figure 4.1 illustrates the number of Ghanaian nationals granted settlement in the United Kingdom and legal permanent residence in the United States between 1975 and 2008 (the US data only cover the time period ranging from 1980-2008). Ghanaian immigration to the United Kingdom started at low levels and increased irregularly throughout the three decades under study. By 2008, the number of Ghanaian immigrants had grown more than sixfold relative to 1975. Low immigration from Ghana in the 1970s reflects the enduring effect of the historic check on immigration from British colonies in Sub-Saharan Africa, notwithstanding the free movement of persons provided by the 1948 British Nationality Act (Abrahmov 2007, Somerville, Sriskandarajah, and Latorre 2009). Notable spikes in the flow of immigrants occurred in 1999, 2003, and 2008, mirroring sudden increases in the overall flow of immigrants to the United Kingdom due to a larger number of students and parents of immigrants, who have been granted settlement[9] (SOPEMI 2001).

As in the case of the United Kingdom, legal emigration from Ghana to the United States started at a low level in 1980 and experienced a twelve-fold increase by 2008. Ghanaian immigration to the United States closely follows the overall trend in legal US immigration. The number of Ghanaian immigrants to the United States experienced an upsurge in 1990, probably reflecting the IRCA induced regularization[10] (Arthur 2012, Hipsman and Meissner 2013), followed by a drop in 1995. Legal Ghanaian immigration to the United States rose sharply again in 1996 as a result of adjustments related to the Section 245(i) of the Immigration Act and of the implementation of the Diversity Visa Program. After a four-year decline, it increased modestly in 2000, due to efforts to reduce the backlog in the adjustment cases.[11] It rapidly increased to reach its historic peak in 2006, just before the reduction in immigration induced by the economic recession.

The figure shows an overall pattern of positive covariation between the Ghanaian immigration flows to the United Kingdom and to the United States, where legal provisions are made to enable migration of close family members under the title of family reunification. This visa category constitutes the single most important contributor to Ghanaian immigration to both the United Kingdom and the United States.

[9] Home Office, Control of Immigration Statistics United Kingdom - various years.

[10] Yearbook of Immigration Statistics, various years.

[11] 2000 Statistical Yearbook of the Immigration and Naturalization Service, U.S. Department of Justice.

Other visa categories, like the employment-based visas, also contributes to the increasing trend of these immigration flows. In the years following 1995, the immigration flows to the United Kingdom and the United States began to diverge as a result of changes in the US immigration policy with the Diversity Visa Program, which came into force in the last quarter of 1994.[12]

In the light of various migration theories, the consistently larger volume of Ghanaian immigration to the United States compared to the United Kingdom could reflect differences in the economic returns to migration. In addition, the socioeconomic differences in emigrants to these two destinations indicate that possible selection mechanisms into each destination might be explained by both the neoclassical economic (human capital) theory and the network theory of migration. To examine these hypotheses, I estimate logistic regressions. Table 4.6 presents the results from a series of logistic regressions of whether emigrants move to the United States with the United Kingdom as the reference category. The first column of the table shows the results of a regression of the variable indicating the United States as the emigration destination on the log-difference between expected earnings in the United States and the United Kingdom, controlling for region of origin. Note that the expected earnings are predicted in the MAFE dataset with parameters from the regressions of migrant earnings on their human capital characteristics (Equation 1) with data from the American Community Survey and the British Labor Force Survey, respectively (See Tables A4.1 and A4.3 in the appendix for the intermediary results from the imputation procedure and earnings regressions). The unadjusted coefficient has a counterintuitive direction, in the sense that it implies that the odds of emigrating to the United States increase as earnings in the United States relative to the United Kingdom decrease. However, the coefficient is not statistically significant. In the

[12] Yearbook of Immigration Statistics, various years.

next two columns, the direction of this association (odds ratios above one) remains unchanged albeit not statistically significant. Results from the full model, which includes demographic and life course characteristics (presented in the last column of the table), indicate that the presence of an additional household member in Europe is associated with higher odds of emigration to the United Kingdom, while a marginal household member in the rest of the world (the Americas, Asia, and Oceania) is correlated with higher odds of emigration to the United States, providing support to the network theory.

4.5. Conclusion

Emigration is an important population process in Ghana and has accelerated rapidly in recent years. Approximately 1.5 million Ghanaians are living abroad, about 0.5 million of whom live in the most developed (OECD) countries (Schans and Bruno Schoumaker 2013). In the United States, Ghanaian immigrants represent the third largest national group among African immigrants (Arthur 2012), after Nigerians and Ethiopians (Zong and Batalova 2014). In recent years, a counter-stream of return migration has been part of the migration experience of an increasing number of Ghanaians (Anarfi, Quartey, and Agyei 2010). However, most studies that have investigated emigration have not paid sufficient attention to return migration, nor have they examined how the choice of migration destinations are indicative of possible selectivity across groups. Furthermore, very few studies have explored plausible mechanisms that work to define observed selectivity among migrants to different destinations in the African context.

The socioeconomic data provided by the MAFE project makes it possible not only to uncover selectivity among current emigrants, return migrants, and non-migrants, but also to investigate such selectivity among emigrants to various destinations. Further information, obtained from the American Community Survey and the British Labor Force Survey, provides the means to explore plausible mechanisms explaining migrant selectivity as put forth by leading theories of migration. Such mechanisms include earnings differentials and networks (in the origin and destination countries).

Consistent with findings from the literature on selectivity, this study finds positive selectivity among Ghanaian emigrants compared to the resident population of Ghana (non-migrants and return migrants). A finer analysis further indicates selectivity among current emigrants compared to both return migrants and non-migrants. While return migrants tend to come from more advantaged background than non-migrants, they are negatively selected based on their own education. The negative selectivity among return migrants compared to non-migrants may stem from cohort differences in educational attainment. Because return migrants tend to be older than non-migrants, with a mean difference in age greater than five years, and because Ghana has experienced education expansion since the 1960s (Akyeampong 2010), younger cohorts would likely average higher levels of education, translating into seemingly negative selectivity among return migrants compared to non-migrants in terms of educational attainment despite the fact that return migrants might have been positively selected compared to their counterpart non-migrants from the same cohort.

Importantly, the nature of socioeconomic selectivity varies depending on whether the education of the household head or the household member's own education is used. For example, while household member's own education indicates a negative selection among return migrants compared to non-migrants, the education of the household head suggests a positive selection among return migrants relative to non-migrants. Thus, these findings highlight the fact that although household member's own education and the education of their household head are positively correlated, they capture distinct socioeconomic dimensions. This might provide an explanation for why some studies on migrant selectivity reach different conclusions.

My findings further show differential selectivity among emigrants to various destinations. They lend support to the neoclassical economic theories, as emigration to countries that are geographically close to Ghana is associated with less selectivity than emigration to more distant countries. This relationship likely stems from variation in immigration policies and travel costs. Although the distance between Ghana and the destination country may be an important factor explaining selectivity, the level of development in the destination countries, a coarse measure of the needs in skills, could well play an important role in determining who moves where.

Finally, my findings are consistent with past research on the role of earnings differentials in explaining the choice of migration destination and on networks. While I find no effect of earnings differentials in explaining the choice of migration destination, migration networks both in destination are important factors that affect the choice of the migration destination.

This study has a number of limitations. First, the lack of information on pre-migration characteristics may affect my results. The measures of educational attainment among emigrants can be affected, since some emigrants may have gained education at destination, before settling down. This raises the concern that such measures may not adequately capture ex-ante selectivity. To the extent that education is predetermined in the late twenties, my findings are indeed exposed to a small bias. Second, although the MAFE project offers socioeconomic information, the relatively small sample size combined with the fact that the sample is not nationally representative of the Ghanaian population limits my ability to conduct detailed analysis for all emigration destinations with a sufficient statistical power and to estimate population-level migration prevalence from these data. Third, there are other socioeconomic characteristics that can further explain the education selectivity among migrants, such as wealth, that are not available in the MAFE dataset. Finally, the proportion of cases with missing information about earnings in the United Kingdom and the United States may well affect my results.

The nature of selectivity among emigrants from poor countries has been an important research topic in the migration literature as it is closely connected to the so-called brain-drain. While recent studies have shown some beneficial effects of emigration mainly through return migration and the flow of remittances, such effects are dependent on the composition of the flow of return migration (Ambrosini et al. 2015, De Vreyer, Gubert, and Robilliard 2010, Docquier and Rapoport 2012). Because my findings suggest a negative selection in return migrants compared to those who remain abroad, they strike some warning notes in the sense that those with more human capital might not be returning. They also highlight that migration theories using a network framework provide more sway in explaining migration. The relatively high rate of return migration prompts the investigation of the duration of the time spent abroad. This suggests a promising avenue for future research on selectivity among emigrants, which should focus on selectivity factors associated with the duration of migratory episodes.

4.6. References

Abrahamova, Natalie. 2007. "Immigration Policy in Britain Since 1962." Masters.

Ackah, C., and D. Medvedev. 2012. "Internal migration in Ghana: determinants and welfare impacts." International Journal of Social Economics, Vol. 39 Iss 10 pp. 39:764 - 784.

Akee, R. 2010. "Who Leaves? Deciphering Immigrant SelfSelection from a Developing Country." Economic Development and Cultural Change 58 (2):pp. 323-344.

Akyeampong, Kwame. 2010. "Educational expansion and access in Ghana: A review of 50 years of challenge and progress."

Ambrosini, J William, Karin Mayr, Giovanni Peri, and Dragos Radu. 2015. "The selection of migrants and returnees in Romania." Economics of Transition 23 (4):753-793.

Anarfi, J., P. Quartey , and J. Agyei. 2010. "Key Determinants of Migration among Health Professionals in Ghana."

Arthur, Mr John. 2012. The African diaspora in the United States and Europe: the Ghanaian experience: Ashgate Publishing, Ltd.

Azam, Jean-Paul, and Flore Gubert. 2006. "Migrants' remittances and the household in Africa: a review of evidence." Journal of African Economies 15 (2):426.

Baizan, P. 2014. "What drives Senegalese migration to Europe? The role of economic restructuring, labor demand and the multiplier effect of networks." DemoSoc Working Paper.

Baizan, P., and A. Gonzalez-Ferrer. 2010. "Destination choices among Senegalese migrants in Europe. The role of labor market and immigration policies."

Beauchemin, C, F Ametepe, A Bring, A Caporali, and T Lejbowicz. 2014. Introduction to the MAFE datasets. MAFE Methodological note 7.

Black, R., A. Gonzalez-Ferrer, E. Kraus, O. Obucina, and P. Quartey. 2013. "Determinants of Migration between Ghana and Europe.

Richard Black, Amparo Gonzlez-Ferrer, Elisabeth Kraus, Ognjen Obucina & Peter Quartey."

Borjas, G. 1994. "The Economics of Immigration." Journal of Economic Literature XXXII:pp. 16671717.

Borjas, George J. 1989. "Immigrant and emigrant earnings: A longitudinal study." Economic inquiry 27 (1):21-37.

Borjas, George J, and Bernt Bratsberg. 1994. Who leaves? The outmigration of the foreign-born. National Bureau of Economic Research.

Capps, Randy, Kristen McCabe, and Michael Fix. 2012. "Diverse streams: African migration to the United States." Migration Policy Institute: Washington, DC.

Castagnone, E., T. Nazio, L. Bartolini, and B. Schoumaker. 2015. "Understanding Transnational Labour Market Trajectories of African-European Migrants: Evidence from the MAFE Survey." International Migration Review 49 (1).

Castles, S., H. De Haas, and M. Miller. 2014. The Age of Migration. 5th Edition ed: The Guilford Press.

Chiswick, B. R. 1978. "The Effect of Americanization on the Earnings of Foreign-Born Men." The Journal of Political Economy:897-922.

Chiswick, Barry R. 1986. "Is the new immigration less skilled than the old?" Journal of Labor Economics 4 (2):168-192.

Constant, Amelie F, and Klaus F Zimmermann. 2003. "The dynamics of repeat migration: a Markov chain analysis."

Constant, Amelie, and Douglas S Massey. 2004. "Self-selection, earnings, and out-migration: A longitudinal study of immigrants to Germany." In How Labor Migrants Fare, 73-95. Springer.

Czaika, Mathias, and Hein Haas. 2016. "The effect of visas on migration processes." International Migration Review.

De Vreyer, Philippe, Flore Gubert, and Anne-Sophie Robilliard. 2010. "Are there returns to migration experience? An empirical analysis using data on return migrants and non-migrants in West Africa." Annals of Economics and Statistics/Annales dconomie et de statistique:307-328.

Docquier, Frdric, and Hillel Rapoport. 2007. "Skilled migration: the perspective of developing countries."

Docquier, Frdric, and Hillel Rapoport. 2012. "Globalization, brain drain, and development." Journal of Economic Literature 50 (3):681-730.

Duleep, Harriet Orcutt. 1994. "Social security and the emigration of immigrants." Soc. Sec. Bull. 57:37.

Dustmann, Christian, Itzhak Fadlon, and Yoram Weiss. 2011. "Return migration, human capital accumulation and the brain drain." Journal of Development Economics 95 (1):58-67.

Dustmann, Christian, and Albrecht Glitz. 2011. "Migration and education." Handbook of the Economics of Education 4:327-439.

Dustmann, Christian, and Oliver Kirchkamp. 2002. "The optimal migration duration and activity choice after re-migration." Journal of development economics 67 (2):351-372.

Dustmann, Christian, and Yoram Weiss. 2007. "Return migration: theory and empirical evidence from the UK." British Journal of Industrial Relations 45 (2):236-256.

Fafchamps, M. , and F. Shilpi. 2013. "Determinants of the Choice of Migration Destination." Oxford Bulletin of Economics and Statistics 75:388409.

Feliciano, Cynthia. 2005. "Educational selectivity in US immigration: How do immigrants compare to those left behind?" Demography 42 (1):131-152.

Flahaux, Marie-Laurence. 2015. "Return Migration to Senegal and the Democratic Republic of Congo: Intention and Realization." Population 70 (1):97-125.

Flahaux, MarieLaurence. 2016. "The role of migration policy changes in Europe for return migration to Senegal." International Migration Review.

Funkhouser, E. 2009. "The Choice of Migration Destination: A Longitudinal Approach using Pre-Migration Outcomes." Review of Development Economics 13 (4):626-640.

Gaul, Patrick. 2014. "Who comes back and when? Return migration decisions of academic scientists." Economics Letters 124 (3):461-464.

Ghana Statistical Service. 2012. Population and Housing Census, Summary Report of Final Results.

Greenwood, M. J. 2005. "Modeling Migration." Encyclopedia of Social Measurement 2.

Harris, J. R., and M. P. Todaro. 1970. "Migration, Unemployment and Development: A Two-Sector Analysis."

Hatton, T. J., and J. G. Williamson. 2001. "Demographic and Economic Pressure on Emigration out of Africa." NBER Working Paper.

Hatton, T. J., and J. G. Williamson. 2002. "What Fundamentals Drive World Migration?" NBER Working Paper.

Hipsman, Faye, and Doris Meissner. 2013. "Immigration in the United States: New economic, social, political landscapes with legislative reform on the horizon." Migration Information Source.

Jasso, Guillermina, Douglas S. Massey, Mark R. Rosenzweig, and James P. Smith. 2004. Immigrant health: selectivity and acculturation. IFS Working Papers, Institute for Fiscal Studies (IFS).

Jasso, Guillermina, and Mark R Rosenzweig. 1988. "How well do US immigrants do?

Vintage effects emigration selectivity and occupational mobility."

Lindstrom, David P, and Douglas S Massey. 1994. "Selective emigration, cohort quality, and models of immigrant assimilation." Social Science Research 23 (4):315-349.

Lucas, R. E. B. 2013. "African Migration." The Handbook on the Economics of International Migration.

Massey, D. 1987. "Understanding Mexican Migration to the United States." American Journal of Sociology 92:1372-1402.

Massey, D. 1999. "Why Does Immigration Occur? A Theoretical Synthesis." Handbook of International Migration, The.

Massey, Douglas S, Joaquin Arango, Graeme Hugo, Ali Kouaouci, Adela Pellegrino, and J Edward Taylor. 1993. "Theories of international migration: A review and appraisal." Population and development review:431-466.

Moretto, M., and S. Vergalli. 2008. "Migration dynamics." Journal of Economics 93:223265.

Reagan, Patricia B, and Randall J Olsen. 2000. "You can go home again: Evidence from longitudinal data." Demography 37 (3):339-350.

Schans, D., and B. Bruno Schoumaker. 2013. "Changing patterns of Ghanaian migration." MAFE Working Paper 20.

Schmelz, A. 2009. The Ghanaian Diaspora in Germany Its Contribution to Development in Ghana. Deutsche Gesellschaft fur Technische Zusammenarbeit.

Schoumaker, B., M-L. Flahaux, and C. Beauchemin. 2015. "Sub-Saharan Migration to Europe in Times of Restriction An empirical test of substitution effects."

Senne, Jean-Noel, Isabelle Chort, and Flore Gubert. 2011. "Migrant Networks as a Basis for Social Control: Remittance Obligations among Senegalese in France and Italy."

Sjaastad, Larry A. 1962. "The costs and returns of human migration." Journal of political Economy 70 (5, Part 2):80-93.

Somerville, Will, Dhananjayan Sriskandarajah, and Maria Latorre. 2009. "United Kingdom: a reluctant country of immigration." Migration Policy Institute.

SOPEMI, OECD. 2001. Trends in International Migration: Annual Report 2001. Paris: Organisation for Economic Co-operation and Development.

Stark, O., and J. E. Taylor. 1991. "Migration Incentives, Migration Types: The Role of Relative Deprivation." The Economic Journal 101 (408):pp. 1163-1178.

Taylor, JE, GF De Jong, JT Fawcett, AR Zolberg, MM Kritz, LL Lim, H Zlotnik, M Weiner, and R Skeldon. 1989. "Differential migration networks information and risk." International Migration Review 23 (3):13-58.

Todaro, Michael. 1980. "Internal migration in developing countries: a survey." In Population and economic change in developing countries, 361-402. University of Chicago Press.

van Dalen, H., G. Groenewold, and J. J. Schoorl. 2003. "Out of Africa: What drives the Pressure to emigrate?".

Wahba, Jackline. 2015. "Selection, selection, selection: the impact of return migration." Journal of Population Economics 28 (3):535-563.

Zong, Jie, and Jeanne Batalova. 2014. "Sub-Saharan African Immigrants in the United States." Migration Information Source: The Online Journal of the Migration Policy Institute.

4.7. Tables

Table 4.1: Sample Characteristics by Emigration Status - Unless Otherwise Stated, Entries are Percentages (MAFE Ghana 2009, ages 25-64)

Measures		All	Non-migrants	Return migrants	Current Emigrants
Relative share		100.0	63.5	10.0	26.6
Education of household head					
	No education	16.1	17.3	15.3	12.7
	Primary/Middle/JHS	44.0	47.2	39.5	35.9
	Secondary/SHS/voc.	21.0	18.9	18.8	27.7
	Higher education	18.9	16.5	26.5	23.7
Respondent's education					
	No education	6.4	5.8	19.4	4.0
	Primary/Middle/JHS	41.8	46.0	28.4	33.5
	Secondary/SHS/voc.	28.7	29.9	20.6	27.6
	Higher education	23.2	18.3	31.6	35.0
Relationship to household head					
	Head	28.3	34.8	58.1	0.0
	Spouse	14.5	16.8	12.8	8.1
	Child	34.5	40.7	24.8	19.0
	Siblings	11.2	3.8	1.7	36.1
	Other rel.	11.5	3.9	2.6	36.8
Mean age at migration					30.18
SD					[0.43]
Mean age		39.1	38.0	43.1	41.0
SD		[0.24]	[0.30]	[0.81]	[0.46]
Age group					
	21-34	38.0	43.9	17.2	26.8
	35-49	44.3	40.1	53.8	53.8
	50 +	17.8	16.1	29.0	19.4
Activity					
	Work	84.3	82.0	82.9	91.3
	Study	4.5	4.6	1.9	5.1
	Unemployed	5.0	6.0	8.9	1.1
	Other activity	6.2	7.4	6.2	2.5
Gender					
	Female	49.9	57.2	30.4	34.4
	Male	50.1	42.8	69.6	65.6
Marital status					
	Unmarried	32.4	36.7	24.1	22.1
	Married	67.6	63.3	75.9	77.9
Year of emigration					
	before 1990	-	-	-	12.8
	1990-1999	-	-	-	36.2
	2000+	-	-	-	51.0
Mean household size		7.2	7.1	6.9	7.4
SD		[0.08]	[0.10]	[0.26]	[0.14]
Mean number of household members in Europe		0.58	0.39	0.52	1.17
		[0.02]	[0.02]	[0.06]	[0.05]
Mean number of household members in Africa		0.12	0.08	0.22	0.21
		[0.01]	[0.01]	[0.06]	[0.03]
Mean number of household members elsewhere		0.57	0.37	0.41	1.20
SD		[0.02]	[0.02]	[0.07]	[0.06]
Region					
	Accra	56.3	56.7	69.0	51.6
	Kumasi	43.7	43.3	31.0	48.4
N		3,837	2,436	382	1,019

Source: Migration from Africa to Europe (MAFE), 2009 Ghana survey
Standard Errors in brackets
Sample characteristics are based on weighted data. The numbers of cases are unweighted

Table 4.2: Sample Characteristics by Emigration Destination - Unless Otherwise Stated, Entries are Percentages (MAFE Ghana 2009, ages 25-64)

Measures	Current	Destinations					
	emigrants	Europe	Americas	Africa	Other dest.[a]	UK	USA
Relative share	100.0	47.0	40.5	9.5	2.9	24.5	34.5
Education of household head							
No education	12.7	10.3	10.2	34.4	18.7	6.0	10.8
Primary/Middle/JHS	35.9	33.0	36.9	46.0	39.3	30.4	37.5
Secondary/SHS/voc.	27.7	31.4	28.6	12.5	6.0	29.6	26.0
Higher education	23.7	25.3	24.3	7.1	36.0	34.0	25.8
Household member's education							
No education	4.0	2.0	0.5	32.2	0.0	1.4	0.6
Primary/Middle/JHS	33.5	35.4	31.5	40.1	12.9	24.4	30.8
Secondary/SHS/voc.	27.6	31.8	27.3	14.2	6.6	34.9	25.8
Higher education	35.0	30.8	40.6	13.5	80.5	39.3	42.8
Relationship to household head							
Head	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Spouse	8.1	7.8	8.5	7.1	12.2	5.4	8.6
Child	19.0	16.9	18.1	21.7	50.8	15.0	19.2
Siblings	36.1	40.2	33.8	33.5	12.4	39.8	34.4
Other rel.	36.8	35.2	39.6	37.7	24.6	39.8	37.7
Mean age at migration	30.2	30.1	30.5	28.0	33.0	30.3	30.7
SD	[0.43]	[0.53]	[0.75]	[1.87]	[1.72]	[0.74]	[0.84]
Mean age	41.0	40.4	41.7	40.7	41.3	40.3	41.8
SD	[0.46]	[0.56]	[0.81]	[1.79]	[2.36]	[0.88]	[0.92]
Age group							
21-34	26.8	27.2	24.5	35.0	27.6	25.8	25.0
35-49	53.8	57.2	53.2	39.0	51.1	56.3	51.7
50 +	19.4	15.5	22.3	26.0	21.3	17.9	23.3
Activity							
Work	91.3	93.6	91.1	84.3	79.9	90.4	90.3
Study	5.1	4.5	4.9	3.3	20.1	6.9	5.4
Unemployed	1.1	1.4	0.6	2.2	0.0	2.0	0.6
Other activity	2.5	0.6	3.4	10.2	0.0	0.7	3.6
Gender							
Female	34.4	30.8	36.8	38.3	46.0	39.6	35.6
Male	65.6	69.2	63.2	61.7	54.0	60.4	64.4
Marital status							
Unmarried	22.1	17.0	25.7	34.1	19.7	20.0	25.6
Married	77.9	83.0	74.3	65.9	80.3	80.0	74.4
Year of emigration							
before 1990	12.8	9.6	13.7	26.4	11.5	9.1	13.2
1990-1999	36.2	35.1	39.9	22.7	44.3	33.4	38.0
2000+	51.0	55.3	46.4	51.0	44.3	57.5	48.7
Mean household size	7.4	7.3	7.5	7.8	6.8	7.9	7.6
SD	[0.14]	[0.20]	[0.20]	[0.48]	[0.97]	[0.31]	[0.22]
Mean number of household members in Europe	1.2	1.9	0.55	0.3	0.4	2.1	0.6
	[0.05]	[0.07]	[0.07]	[0.09]	[0.15]	[0.11]	[0.08]
Mean number of household members in Africa	0.2	0.1	0.1	1.7	0.0	0.0	0.1
	[0.03]	[0.02]	[0.03]	[0.13]	[0.01]	[0.01]	[0.03]
Mean number of household members elsewhere	1.2	0.5	2.1	0.3	1.7	0.6	2.1
SD	[0.06]	[0.07]	[0.09]	[0.14]	[0.27]	[0.10]	[0.09]
Region							
Accra	51.6	47.7	56.1	48.1	61.0	56.4	57.0
Kumasi	48.4	52.3	43.9	51.9	39.0	43.6	43.0
N	1,019	479	413	97	30	250	352

Source: Migration from Africa to Europe (MAFE), 2009 Ghana survey
Note: ([a]) Other destinations include: Oceania and Asia
Standard Errors in brackets
Sample characteristics are based on weighted data. The numbers of cases are unweighted

Table 4.3: Odds Ratios from Logistic Regression of Whether Household Member is Currently Living Abroad among Ghanaians ages 25-64, Omitted Category = Non-Migrants+Return Migrants (2009 Ghana MAFE)

Measures	Model 1	Model 2	Model 3
Household member's education			
[No education]			
Primary/Middle/JHS	1.38	1.29	1.27
Secondary/SHS/voc.	1.86**	1.40	1.10
Higher education	3.66***	3.09***	3.04***
Age group			
[35-49]			
21-34	0.45***	0.43***	0.47***
50 +	0.87	0.84	0.94
Education of household head			
[No education]			
Primary/Middle/JHS		1.01	1.52
Secondary/SHS/voc.		1.91***	1.92*
Higher education		1.22	1.03
Gender			
[Female]			
Male			1.91***
Marital status			
[Unmarried]			
Married			2.69***
Relationship to household head			
[Head+Siblings]			
Spouse			0.83
Child			1.11
Other rel.			16.42***
Log-household size			0.29***
Number of household members in Europe			2.61***
Number of household members in Africa			3.26***
Number of household members elsewhere			2.43***
Region			
[Kumasi]			
Accra			0.72**
N	3,837	3,837	3,837

Source: Migration from Africa to Europe (MAFE), 2009 Ghana survey
Note: Estimates are weighted. Reference group for each variable is given in brackets.
*** p<0.01, ** p<0.05, * p<0.1

Table 4.4: Odds Ratios from Multinomial Logit Regression of Emigration Status among Ghanaians ages 25-64, Omitted Category = Non-Migrants (2009 Ghana MAFE)

Measures	Model 1		Model 2		Model 3	
	Curr. Emig.	Ret. Mig.	Curr. Emig.	Ret. Mig.	Curr. Emig.	Ret. Mig.
Household member's education						
[No education]						
Primary/Middle/JHS	1.06	0.19***	0.95	0.12***	0.83	0.12***
Secondary/SHS/voc.	1.47	0.24***	1.04	0.14***	0.72	0.12***
Higher education	3.26***	0.66	2.54**	0.34**	2.25*	0.28**
Age group						
[35-49]						
21-34	0.40***	0.26***	0.38***	0.25***	0.41***	0.27***
50 +	0.90	1.25	0.86	1.19	0.97	1.14
Education of household head						
[No education]						
Primary/Middle/JHS			1.08	2.06**	1.68	2.28**
Secondary/SHS/voc.			2.09***	2.70***	2.15**	2.61**
Higher education			1.38	3.09***	1.19	2.89**
Gender						
[Female]						
Male					2.19***	2.87***
Marital status						
[Unmarried]						
Married					2.77***	1.25
Relationship to household head						
[Head+Siblings]						
Spouse					0.79	0.74
Child					1.07	0.83
Other rel.					15.72***	0.73
Log-household size					0.28***	0.74
Number of household members in Europe					2.73***	1.40***
Number of household members in Africa					3.61***	1.87***
Number of household members elsewhere					2.46***	1.14
Region						
[Kumasi]						
Accra					0.76*	1.49*
N	3,837	3,837	3,837	3,837	3,837	3,837

Source: Migration from Africa to Europe (MAFE), 2009 Ghana survey
Note: Estimates are weighted. Reference group for each variable is given in brackets.
*** p<0.01, ** p<0.05, * p<0.1

Table 4.5: Odds Ratios from Multinomial Logit Regression of Emigration Status among Ghanaians ages 25-64, Omitted Category = Emigrants in Europe (2009 Ghana MAFE)

Measures	Model 4a		Model 4b	
	Americas	Africa	Americas	Africa
Household member's education				
[No education]				
Primary/Middle/JHS	4.03	0.43		
Secondary/SHS/voc.	6.28	0.06**		
Higher education	7.33	0.03***		
Age group				
[35-49]				
21-34	0.72	1.12	0.74	1.15
50 +	1.11	3.58*	1.07	2.3
Education of household head				
[No education]				
Primary/Middle/JHS			1.36	0.62
Secondary/SHS/voc.			1.01	0.28
Higher education			0.86	0.14**
Gender				
[Female]				
Male	0.61	0.22	0.66	0.27
Marital status				
[Unmarried]				
Married	0.57	0.21*	0.58	0.16**
Relationship to household head				
[Head+Siblings]				
Spouse	1.5	7.95**	1.36	6.20**
Child	0.98	0.59	1.14	0.36
Other rel.	1.64	1.33	1.76	1.63
Log-household size	1.05	0.5	0.97	0.47
Number of household members in Europe	0.19***	0.11***	0.20***	0.09***
Number of household members in Africa	0.83	103.48***	0.7	54.69***
Number of household members elsewhere	5.15***	0.54	5.17***	0.48
Year of emigration				
[2000+]				
before 1990	3.04**	4.03**	2.73**	8.11***
1990-1999	1.04	0.74	0.96	1.28
Region				
[Kumasi]				
Accra	0.60*	4.42**	0.79	1.98
N[a]	989	989	989	989

Source: Migration from Africa to Europe (MAFE), 2009 Ghana survey
([a]) Sample excludes Emigrants to Asia and Oceania
Note: Estimates are weighted. Reference group for each variable is given in brackets.
*** p<0.01, ** p<0.05, * p<0.1

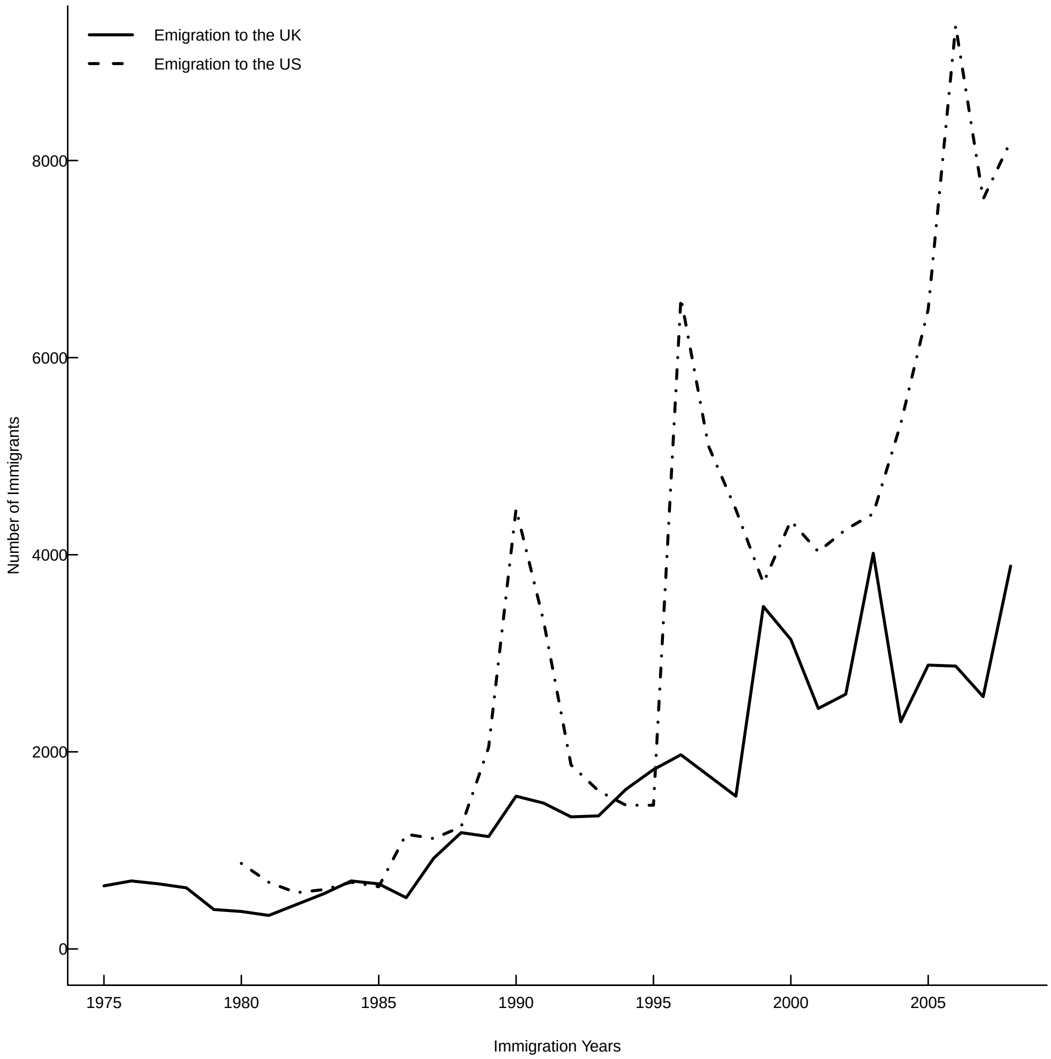
Table 4.6: Odds Ratios from Logistic Regression of Whether Emigrants Move to the United States among Emigrants 25-64, (2009 Ghana MAFE)

Measures	Dest. (US=1)	Dest. (US=1)	Dest. (US=1)
Log - Earnings difference between the US and UK	0.50	0.35	0.20
Region of origin			
Kumasi			
G. Accra	1.02	0.99	0.60
Year of emigration			
[2000+]			
before 1990		1.27	3.77**
1990-1999		0.71	0.89
Gender			
[Female]			
Male			1.00
Marital status			
[Unmarried]			
Married			1.07
log-household size			0.65
Number of household members in Europe			0.23***
Number of household members in Africa			1.62
Number of household members elsewhere			5.93***
N	602	602	602

Source: Migration from Africa to Europe (MAFE), 2009 Ghana survey
Note: Estimates are weighted. Reference group for each variable is given in brackets.
*** p<0.01, ** p<0.05, * p<0.1

4.8. Figures

Figure 4.1: Legal Emigration from Ghana to the United Kingdom and United States, 1975 - 2008



Source: Office of Immigration Statistics (UK) and Home Office (UK)

4.9. Appendix

Table A4.1: Distribution of Missing Values by Variable and Survey

Panel A: American Community Survey

Variables		Missing	Non-missing	Unique Values
Log-earnings		2,794	3,029	>500
Age		0	5,823	-
Male		0	5,823	-
Household Member's education				
	High school Ref.	0	5,823	-
	No education	0	5,823	-
	< High school	0	5,823	-
	College +	0	5,823	-

Panel B: Labor Force Survey

Variables		Missing	Non-missing	Unique Values
Log-earnings		642	167	165
Age		0	809	-
Male		0	809	-
Household Member's education				
	High school Ref.	0	809	-
	No education	0	809	-
	< High school	0	809	-
	College +	0	809	-

Source: American Community Survey (US), waves 2005-2010
Labor Force Survey (UK), 2010-2012

Table A42: Imputation Diagnostics by Survey (Number of Imputations = 50)

Panel A: American Community Survey

Variables	Imputation Variance			RVI	FMI	Relative efficiency
	Within	Between	Total			
Age	0.000	0.000	0.000	0.910	0.481	0.990
Age-squared	0.000	0.000	0.000	0.915	0.483	0.990
Male	0.000	0.000	0.001	1.214	0.554	0.989
Household Member's education						
High school Ref.						
No education	0.007	0.011	0.019	1.610	0.623	0.988
< High school	0.001	0.001	0.002	1.313	0.573	0.989
College +	0.001	0.002	0.002	2.218	0.695	0.986
Constant	0.032	0.028	0.061	0.883	0.474	0.991

Panel B: Labor Force Survey

Variables	Imputation Variance			RVI	FMI	Relative efficiency
	Within	Between	Total			
Age	0.000	0.001	0.002	2.896	0.751	0.985
Age-squared	0.000	0.000	0.000	2.966	0.755	0.985
Male	0.002	0.007	0.009	3.699	0.794	0.984
Household Member's education						
High school Ref.						
No education	0.010	0.036	0.046	3.841	0.801	0.984
< High school	0.006	0.020	0.026	3.593	0.790	0.984
College +	0.005	0.018	0.023	3.681	0.794	0.984
Constant	0.235	0.588	0.835	2.547	0.726	0.986

Source: American Community Survey (US), waves 2005-2010
Labor Force Survey (UK), 2010-2012

Table A4.3: Regression of Migrant Earnings in the UK and the US on their Human Capital Characteristics - Ages 25-64

Measures	ln (UK earn.)	ln (US earn.)
Age	-0.02	0.06***
	[0.05]	[0.01]
Age squared	0.00	-0.00***
	[0.00]	[0.00]
Emigrant is male	0.11	0.20***
	[0.08]	[0.03]
Emigrant education		
High school Ref.	ref.	ref.
No education	-0.07	-0.33**
	[0.20]	[0.14]
< High school	0.19	0.05
	[0.17]	[0.04]
College +	0.29**	0.24***
	[0.15]	[0.05]
Constant	1.64	0.17
	[1.02]	[0.25]
N	809	5,823

Source: American Community Survey (US), waves 2005-2010
Labor Force Survey (UK), 2010-2012
Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1