SANDIA REPORT

SAND2015-10235
Unlimited Release
November 2015

Climate Induced Spillover and Implications for U.S. Security

Vincent C. Tidwell, Asmeret Naugle, George A. Backus, Kathryn Lott, Elizabeth Kistin-Keller, Prabuddha Sanyal, Peter Kobos, Daniel Villa

Prepared by Vincent Tidwell
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@osti.gov
Online ordering: http://www.osti.gov/scitech

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.gov
Online order: http://www.ntis.gov/search

SAND2015-10235
Unlimited Release
November 2015

Climate Induced Spillover and Implications for U.S. Security

Vincent C. Tidwell, Asmeret Naugle, George A. Backus, Kathryn Lott, Elizabeth Kistin-Keller,
Prabuddha Sanyal, Peter Kobos, Daniel Villa
Earth Systems Analysis
Cognitive Systems
Systems Research and Analysis
Sandia National Laboratories
P.O. Box 5800
Albuquerque, New Mexico 87185-MS1137

Abstract

Developing nations incur a greater risk to climate change than the developed world due to poorly managed human/natural resources, unreliable infrastructure and brittle governing/economic institutions. These vulnerabilities often give rise to a climate induced "domino effect" of reduced natural resource production-leading to economic hardship, social unrest, and humanitarian crises. Integral to this cascading set of events is increased human migration, leading to the "spillover" of impacts to adjoining areas with even broader impact on global markets and security. Given the complexity of factors influencing human migration and the resultant spill-over effect, quantitative tools are needed to aid policy analysis. Toward this need, a series of migration models were developed along with a system dynamics model of the spillover effect. The migration decision models were structured according to two interacting paths, one that captured long-term "chronic" impacts related to protracted deteriorating quality of life and a second focused on short-term "acute" impacts of disaster and/or conflict. Chronic migration dynamics were modeled for two different cases; one that looked only at emigration but at a national level for the entire world; and a second that looked at both emigration and immigration but focused on a single nation. Model parameterization for each of the migration models was accomplished through regression analysis using decadal data spanning the period 1960-2010. A similar approach was taken with acute migration dynamics except regression analysis utilized annual data sets limited to a shorter time horizon (2001-2013). The system dynamics spillover model was organized around two broad modules, one simulating the decision dynamics of migration and a second module that treats the changing environmental conditions that influence the migration decision. The environmental module informs the migration decision, endogenously simulating interactions/changes in the economy, labor, population, conflict, water, and food. A regional model focused on Mali in western Africa was used as a test case to demonstrate the efficacy of the model.

CONTENTS

1. INTRODUCTION

1.1. Background

Some 215 million people or 3 percent of the world's population are believed to live outside their countries of birth (United Nations 2009a), while millions more have migrated internally (United Nations 2009b). Migration is often undertaken to improve one's quality of life, in response to such factors as the economy/employment (Massey et al. 1993), land degradation (Ghimire and Mohai 2005), social networks (Davis et al. 2013), and community factors (Grote et al. 2006). In other cases migration is forced, driven by conflict (IDMC 2011; Salehyn and Gleditsch 2006), environmental disasters (El-Hinnawi 1985), or other influences. Migration is a decision that impacts the welfare of the household, the home community, and in the end the whole economy (Azam and Gubert 2006). Spillover effects of increased migration may include greater need for costly humanitarian aid, increased crime and terrorism, interruption of international trade, and mobilization of peace keeping forces (CAN 2009; Defense Science Board 2011; National Intelligence Council 2012).

The growing concern over climate change has drawn recent attention to human migration, suggesting that intensifying floods, droughts, and sea level rise could result in unprecedented migration (Myers 2002). Projections of environmentally induced migration vary widely, from 200 million (Brown 2008) to 700 million (Christian Aid 2007) by 2050. While many may question these numbers (Black 1998; 2001), few disagree with the fact that related environmental challenges will put increased stress on at-risk populations and may motivate internal and international migration.

Homer-Dixon (1991; 1994) gives case study evidence that links population growth, environmental deterioration and political violence to migration. Reuveny and Moore (2009) found deteriorating environmental conditions in a developing country promotes out-migration to the developed world, all else being equal. In a study of Nepal's Chitwan Valley during the late 1990s environmental deterioration (e.g., declining land cover, increasing population density, perceived declines in agricultural productivity) was found to lead to short-distance moves within the immediate vicinity (Massey et al. 2007). Similarly, analysis of survey data for both migrants and non-migrants in 12 countries suggest that while long-term environmental events, such as droughts, have no significant effect on internal migration, sudden-onset environmental events in the form of floods significantly increase the likelihood of migration. Furthermore, individual perceptions of negative environmental conditions can motivate people to move. They also found that people tend to respond to long-term environmental problems, such as environmental degradation, with adaptation, rather than migration. Ultimately, migration dynamics are complicated with environmental conditions being one of many factors that mutually influence migration (Wood 2001).

Analytical tools are needed to assist in identifying populations at greatest risk and exploring robust policy strategies and adaptive measures (UN Framework Convention on Climate Change 2013). A variety of approaches have been taken toward analyzing environmental impacts on migration trends. Perch-Nielsen (2004) argued that migration induced by climatic hazards has not been integrated into migration models; therefore, existing climate and migration models cannot be simply linked. She proposed four conceptual models linking climate change and migration, addressing sea level rise, floods, tropical cyclones and drought. Other qualitative/conceptual models have been developed as the result of a variety of environmental induced migration case studies (McLeman and Smit 2006; Gilbert and McLeman 2010; Black et al. 2011). Empirically based models have been developed through regression of large sets of data generally structured around a conceptual model of migration; for example, informal cost-benefit (Reuveny and Moore 2009) and the gravity model (Afifi and Warner 2008). Economic modeling has also been pursued; specifically, a general equilibrium (microeconomic) model of environmental migration (Siyaranamual 2009; Chichilnisky and DiMatteo 1998). Agent based modeling provides a framework to simulate human behavior, allowing agents (individuals or groups of individuals) to interact with the environment in complex ways including environmentally motivated migration. Examples include Mena et al. (2011) that developed an agent-based model to simulate deforestation change associated with land use patterns of frontier migrant framers in Northern Ecuadorian Amazon. Kniveton et al. (2012) used an agent-based model developed around the theory of planned behavior to explore how climate and demographic change combine to influence migration within and from Burkina Faso.

1.2. Motivation and Objectives

Human migration accompanied by the intensifying effects of climate change is a potential national security issue. Developing nations incur a greater risk to climate change than the developed world due to poorly managed human/natural resources, unreliable infrastructure and brittle governing/economic institutions. These vulnerabilities often give rise to a climate induced “domino effect” of reduced natural resource production-leading to economic hardship, social unrest, and humanitarian crises. Integral to this cascading set of events is increased human migration, leading to the “spillover” of impacts to adjoining areas with even broader impact on global markets and security. Given the complexity of factors influencing human migration and the resultant spill-over effect, quantitative tools are needed to aid policy analysis.

Toward this problem a model of climate induced spillover was developed. The unique aspect of this work is the integration of social, economic, infrastructure and resource dynamics/constraints in the context of climate change to provide a comprehensive assessment of their interdependent influence on human migration. The model is also designed to explore alternative future adaptation pathways to understand the efficacy and robustness of alternative policy strategies, determining what pre-emptive adaptive measures are most necessary when and where.

1.3. Approach

The ultimate focus of this work is on our changing climate, its influence on resource provisioning and the resultant threat to the sustainability/stability of the human welfare and security. That is, spillover occurs where climate impacts (e.g., drought, flood, storm) exceed the adaptive capacity of society. The adaptive capacity depends, at least in part, on the diversity, substitutability, redundancy, and resiliency of the impacted society. These measures apply not only to natural resource available but also to the critical infrastructure necessary to abstract/convey the resources, the economic ability to finance/consume/market the resources, and the societal capacity to govern/regulate/manage the natural resources. Communities most vulnerable to spillover are, for example, those that lack a diverse economy, lack the ability to substitute production of one resource to overcome losses in another, lack redundancy in the modes of transportation for key goods to market, and/or lack a resilient governance structure to promote recovery in times of emergency.

A series of migration models were developed along with a system dynamics model of the spillover effect. The migration decision models were structured according to two interacting paths, one that captured long-term "chronic" impacts related to protracted deteriorating quality of life and a second focused on short-term "acute" impacts of disaster and/or conflict. Chronic migration dynamics were modeled for two different cases; one that looked only at emigration but at a national level for the entire world; and a second that looked at both emigration and immigration but focused on a single nation. Model parameterization for each of the migration models was accomplished through regression analysis using decadal data spanning the period 1960-2010. A similar approach was taken with acute migration dynamics except regression analysis utilized annual data sets limited to a shorter time horizon (2001-2013). The system dynamics spillover model was organized around two broad modules, one simulating the decision dynamics of migration and a second module that treats the changing environmental conditions that influence the migration decision. The environmental module informs the migration decision, endogenously simulating interactions/changes in the economy, labor, population, conflict, water, and food. A regional model focused on Mali in western Africa was used as a test case to demonstrate the efficacy of the model.

The models and their application are discussed in the following chapters of the report. These chapters are organized according to six largely stand-alone papers which were prepared on complimentary aspects of the modeling. The relations between these chapters are as follows:

<u>Chapter 2: Modeling International Emigration and Climate Change Effects:</u> This chapter describes efforts toward developing a model of human emigration capturing long-term, chronic dynamics. The model simulates international emigration for 166 countries given projected changes in a variety of measures of human security and adaptive capacity. The chronic migration dynamics modeled here helped inform and structure the coupled system dynamics model described in Chapter 4. This chapter was originally prepared for the 8th Conference on Sustainable Development of Energy, Water, and Environment Systems (SDEWES) held in Dubrovnik, Croatia on September 22-27, 2013.

<u>Chapter 3: Long Term Migration Dynamics:</u> This chapter presents an alternative approach to migration model in Chapter 2 that explicitly addresses both the push and pull migration dynamics (factors influencing both the decision to leave and the decision of where to go). The push-pull migration analysis focuses on a single country, Mali (distinguished by urban and rural populations) and the choice to migrate to neighboring countries (in aggregate), the US or the Rest of World. This mathematical construct was used in the spillover model (Chapter 6).

<u>Chapter 4: Short-Term Migration Dynamics:</u> This chapter documents a model of human migration that responds to the acute or short-term triggers. Short-term migration is

dominantly a push process driven largely by presumed temporary conditions within the home locale, such as military activity, local violence, or natural disaster. This analysis considered short-term migration dynamics for the continent of Africa.

Chapter 5: Spillover Violence: This chapter addresses a related issue to migration, violence. The analysis considers the relative level of violence in a neighboring country as it affects violence in the country of interest. This analysis addresses the potential for the spillover of violence in Mali to neighboring countries and vice versa. This mathematical construct was used in the spillover model (Chapter 6).

Chapter 6: A Regional Model of Human Migration and Climate Change Effects. This chapter documents the entire coupled system dynamics model with specific application to the West African nation of Mali. This test case provides an example of the capabilities of the model and the type of questions that it can be used to address. This chapter was originally prepared for the 2015 Annual Meeting of the International Migration, Integration and Social Cohesion in Europe (IMISOCE) held in Geneva, Switzerland on June 25-27, 2015.

In addition several appendices are included. These appendices provide additional detailed information on the various modeling components discussed in the body of the report.

1.4. References

Afifi, T. and Warner, K., 2008. *The Impact of Environmental Degradation on Migration Flows Across Countries, United Nations Working Paper No. 5/2008*, Bonn.

Azam, J.P., & F. Gubert. 2006. Migrants' Remittances and the Household in Africa: A Review of the Evidence. *Journal of African Economies,* Vol. 15, AERC Supplement 2: 426-462.

Black, R. 1998. *Refugees, Environment and Development*. London: Longman. Black, R. 2001.

Black, R., 2001. Environmental Refugees: Myth or Reality? New Issues in Refugee Research Working Paper 34. Geneva: UN High Commissioner for Refugees

Black, R.B., Adger, N., Arnell, N., Dercon, S., Geddes, A., Thomas, D., 2011. *Foresight: Migration and Global Environmental Change, Final Project Report*, The Government Office for Science, London.

Brown, O., 2008. *Migration and Climate Change, International Organization for Migration (IOM): Research Series No. 31*, IOM Geneva.

Chichilnisky, G. and DiMatteo, M, 1998. *Trade, Migration and Environment: A General Equilibrium Analysis*, G. Chichilnisky, G. Heal, and A Vercelli (eds). Kluwer Academic Publishers.

Christian Aid, 2007. *Human Tide: The real migration crisis,* Christian Aid Report May 2007. London.

Davis, K.F., D'Odorico, P., Laio, F., Ridolfi, L., 2013. Global Spatio-Temporal Patterns in Human Migration: A Complex Network Perspective, *PLoS ONE*, Vol. 8, No. 1.

Defence Science Board, 2011. *Trends and Implication of Climate Change for National and International Security*.

El-Hinnawi, E., 1985. *Environmental Refugees*. United Nations Environment Programme, Nairobi.

Grote, U., Engel, S., Schraven, B., 2006. *Migration due to the Tsunami in Sri Lanka –Analyzing Vulnerability and Migration at the Household Level*. ZEF-Discussion Paper on Development Policy No. 106, Centre for Development Research (ZEF), Bonn, April 2006.

Ghimire, D.J. and Mohai, P.L., 2005. Environmentalism and Contraceptive Use: How People in Less Developed Settings Approach Environmental Issues, *Population and Environment*, Vol. 27, No. 1: 29-61.

Gilbert, G., and McLeman, R., 2010. Household access to capital and its effects on drought adaptation and migration: A case study of rural Alberta in the 1930s. *Population and Environment*, Vol.32, No.1: 3-26.

Homer-Dixon, Thomas F. 1991. On the Threshold: Environmental Changes as Causes of Acute Conflict. *International Security* 16: 76-116.

Intergovernmental Panel on Climate Change (IPCC). 2000. *Emissions Scenarios*. Ed. Nebojsa Nakicenovic and Rob Swart, Cambridge University Press, Cambridge, England. http://www.ipcc.ch/ipccreports/sres/emission/

Kniveton, D.R., Smith, C.D., Black, R., 2012. Emerging Migration Flows in a Changing Climate in Dryland Africa, Nature Climate Change Letters.

Koubi, V., Spilker, G., Schaffer, L. M. and Bernauer, T., 2012. Environmental Degradation and Migration. Available at SSRN: http://ssrn.com/abstract=2107133 or http://dx.doi.org/10.2139/ssrn.2107133

Internal Displacement Monitoring Center and the Norwegian Refugee Council (IDMC), 2011. *Internal Displacement: Global Overview of Trends and Developments in 2010*.

National Intelligence Council, 2012. *Global Trends 2030: Alternative Worlds*. NIC 2012-001.

Massey, D.S. Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., Taylor, J.E., 1993. Theories of Migration: A Review and Appraisal, *Populating and Development Review*, Vol. 19, No. 3: 431-466.

Myers, N., 2002. Environmental Refugees: A Growing Phenomenon of the 21st Century, *Philosophical Transactions of The Royal Society B*. Vol. 357: 609-613.

Massey, D., Axinn, W., and Ghimire, D., 2007. *Environmental Change and Out-Migration: Evidence from Nepal*, Report 07-615, Population Studies Center, University of Michigan, 28p.

McLeman, R., and Smit, B., 2006. Migration as an Adaptation to Climate Change. *Climatic Change*, Vol. 76, No.1-2: 31-53.

Mena, C.F., Walsh, S.J., Frizzelle, B.G., Yao, X.Z., and Malanson, G.P., 2011. Land use change on household farms in the Ecuadorian Amazon: Design and implementation of an agent-based model, Applied Geography, 31(1): 210-222.

Perch-Nielsen, S.L., 2004. *Understanding the effect of climate change on human migration: The contribution of mathematical and conceptual models, diploma thesis*. Swiss Federal Institute of Technology, Zurich.

Reuveny, R. and Moore, W.H., 2009. Does environmental degradation influence migration? Emigration to developed countries in the last 1980 and 1990s, *Social Science Quarterly*, 90(3): 461-479.

Salehyan, I., and Gleditsch, K.S., Refugees and the Spread of Civil War, 2006. *International Organization,* Vol. 60: 335–366.

Siyaranamual, M.D., 2009. *Constructing the Model of Environmental Migration*, Working Paper in Economic and Development Studies No. 200901, Padjadjaran University, 25p.

United Nations, 2009a. *International Migrant Stock: The 2008 Revision*. United Nations Population Division (http://esa.un.org)

United Nations. 2009b. *Monitoring Global Population Trends*. United Nations Population Division (http://www.un.org/en/development/desa/population/publications/urbanization/urban-rural.shtml)

Wood, W.B., 2001. *Ecomigration: Linkages Between Environmental Change and Migration*. Pp. 42-61 in Aristide R. Zolberg and Peter Benda, eds., Global Migrants, Global Refugees. New York and Oxford: Berghahn.

2. MODELING THE EFFECTS OF ENVIRONMENTAL CHANGE ON INTERNATIONAL EMIGRATION

2.1. Abstract

Climate change poses a credible threat to vulnerable populations through intensifying damage to homes and critical infrastructure, reduced food production, compromised health and hygiene, as well as land and environmental degradation. These events combined with other physical and social factors can spiral into economic loss, reduced confidence in government institutions, and conflict–often increasing pressure on human migration. Given the complexity of factors influencing human migration, quantitative tools are needed to aid policy analysis. Toward this need, a dynamic model of human migration is developed structured on the Theory of Planned Behavior. The model simulates international emigration given changes in various measures of the physical and social conditions of humans. Model parameterization was accomplished through regression analysis following a Partial Adjustment Model process using decadal data spanning the period 1960-2010. Analyses were conducted for three cases 1) all nations (166 nations), 2) developing nations, and 3) nations on the continent of Asia. In efforts to validate the all nations emigration model a state-space model was formulated to independently estimate emigration flows at mid-decadal intervals. The state space approach offers a computationally more efficient alternative to fit single level regression models with a large number of independent variables for each country in a panel data setting.

2.2. Introduction

Migration has been a part of the human endeavor since the beginning of time as evidenced by the broad distribution of peoples throughout the world. In the most simple of terms, people migrate in hopes of improving their quality of life (Ravenstein 1889); however, the specific reasons for migration are many and varied. Several causal linkages have been suggested, including conflict (Internal Displacement Monitoring Center 2011; Salehyan and Gleditsch 2006), economy/employment (Massey et al. 1993), land degradation (Ghimire and Mohai 2005), social networks (Davis et al. 2013), community factors (Grote et al. 2006), and environmental disasters (El-Hinnawi 1985). Rarely is there a single causal factor, but rather multiple stressors acting together.

The specter of climate change has drawn recent attention to human migration, suggesting that intensifying floods, droughts, and sea level rise could result in unprecedented migration (Myers2002). Projections of environmentally induced migration vary widely, from 200 million (Brown 2008) to 700 million (Christian Aid 2007) by 2050. While many may question these numbers, few disagree with the fact that such environmental challenges will put increased stress on at-risk populations to migrate. Migration not only impacts the receiving destination, but can spillover internationally through the need for costly humanitarian aid, increased crime and terrorism, interruption of international trade, and mobilization of peace keeping forces (CAN 2009:Defence Science Board 2011; National Intelligence Council 2012). Quantitative tools are needed to assist in identifying populations at greatest risk and exploring robust policy strategies and adaptive measures (United Nations Framework Convention on Climate Change 2013).

The conceptual basis for much of today's migration modeling can be traced back to Ravenstein's seven laws of migration (Ravenstein 1889). Lee (1966) further shaped the conceptual model of migration by recognizing that certain forces tend to push potential migrant from their place of origin while other forces tend to pull the migrant to their point of destination. This "push-pull" process was further envisioned as being modulated by intervening obstacles, which could be physical, social, or political in nature. Additional details have been contributed to the conceptual/quantitative modeling of human migration, elucidating the drivers of migration (i.e., social, political, demographic, economic and environmental), incorporating ideas of adaptive capacity and competing adaptive strategies (i.e., adapt in place, migration) (McLeman and Smit 2006; Black et al. 2011; Gilbert and McLeman 2010).

Moving from conceptual to quantitative modeling poses its challenges as it is difficult both from the complexity of the individual's decision making process and the limited availability of data to establish such dynamics. Empirical methods have found particular value in this space. A common empirical approach involves use of the gravity model, which treats migration as proportional to the population masses at the point of origin and destination and inverse to the distance between the two locations (Faist 2000). Afifi and Warner (2008)] expanded the gravity model to include economic, political, social, and historical/cultural factors while evaluating their statistical significance through regression analysis. Various regression models have been used to explore the impact of environmental decline on emigration Reuveny and Moore (2009), and processes driving Mexico-U.S. migration (Massey and Espinosa 1997).

Numerical models that combine Agent-Based Models (ABMs) with models of the environment in which the agents operate are finding growing application in the analysis of human migration. The ABM simulates the decision dynamics of various agents which can represent individuals, households, or communities. The environmental model encompasses changes in the economy, health, critical infrastructure, climate or any other factor effecting an agent's decision. In this way, agents influence and are influenced by the environment around them as well as other agents they encounter in a simulation. A few examples of this coupled modeling include; migration-land use dynamics in Northeastern Thailand (Walsh et al. 2013); the economic, social and environmental drivers of migration in two cities in the United Kingdom (Zhang and Jager 2011); and climate change impact on extreme poverty, socioeconomic vulnerability and demography leading to increased migration in Bangladesh (Hassani-Mahmooei and Parris 2012). The challenge is identifying the factors that influence the migration decision, accurately representing those factors in the model, and then capturing the appropriate dynamics between the disparate factors (e.g., changes in stream flow impact fisheries production which both intensify the migration decision) and the actions of the agents. Given the heavy data requirements, such models are generally limited to regional analyses.

The objective of this study is to develop a global model of international emigration. The theoretical underpinning for the model is the Theory of Planned Behavior (Ajzen 1991). This national-level model integrates dynamics of vulnerability, adaptive capacity, and social interactions that inform the migration decision. This emigration model is designed to operate within the broader context of integrated assessment modeling of global climate change (Kniveton 2011). The ultimate goal being a framework for assessing how climate change influences such factors as water availability, food production, economic growth, human health/disease, and conflict, which in turn impact human emigration.

2.3. Emigration Model

The present work represents the first step in a broader effort to model human migration. Here the focus is limited to the case of international emigration. Future work will attempt to model the decision dynamics around a migrant's choice of destination. Below a description of the model is provided, including the theoretical underpinnings, data selection and regression analysis.

2.3.1. Theory

The decision to migrate is generally prompted by some change in environment. Vulnerability to the change determines whether adaptive action must be taken. The choice to migration among other options is likely to involve such factors as the perception of the sustainability of living conditions; assumptions about how the event will evolve based on past experience; adaptive capacity both to remain in place (e.g., government support, family/community support) or to migrate (e.g., savings, social networks); and the advice as well as the approval of significant others. The Theory of Planned Behavior (Ajzen 1991) provides a cognitive framework for modeling an agent's migration decision process.

The Theory of Planned Behavior posits that the proximal cause of behavior is 'behavioral intention', a conscious decision to engage in certain behavior. Behavioral intention is modeled as the interaction of behavioral attitude, subjective norms and perceived behavioral control. Behavioral attitudes measure one's perceived betterment by taking an intended action, such as improved wages, living conditions, or personal security. Subjective norms address the influence of networked peer approval on an intended action. Perceived behavioral controls are simply the perception of whether or not one has the assets/experience necessary to undertake an intended action. Kniveton and others (2011; 2012); and De Jong (1999; 2000) have adapted and applied the Theory of Planned Behavior to migration decision-making.

The behavior intent is simply the weighted sum of various measures of behavioral attitude, subjective norms and perceived behavioral control that an agent considers in their migration decision. The challenge is to establish the appropriate weights and measures. For purposes of this study, regression analysis using historical migration data serves as the vehicle for model parameterization.

2.3.2. Model Terms

The intent of this study is limited to modeling international emigration at the national level; that is, simulating emigration dynamics for each nation in the world (designed for application with integrated assessment models for global climate change). Even within this relatively narrow scope the expression of migration can be manifest in different ways depending on the timing, duration, spatial scope and purposefulness of action (Smit et al. 2000; Smit and Wandel 2006). That is, migration can be anticipatory or proactive, short term or long term, intentioned or forced, individual or family. Even the community demographics such as age, gender, education, marital status, and ethnicity play an important role in the potential migrant's decision (McLeman and Smit 2006).

Unfortunately, data (at the global scale) are largely lacking to disaggregate the migrant population into these various communities. Our analysis works to the strength of the available data—utilizing the bilateral migrant stock data available on 10-year increments from 1960-2010 for 232 countries from the World Bank (2011). The bilateral data were aggregated to yield total migrant stock by country and decade. The data were further normalized by the corresponding population in the country of origin, to yield migrants per capita. Finally, these data were referenced to the year 1960 to focus the analysis on the trends in migration.

While the migrant stock data were taken as the dependent variable in the regression analysis, associated independent variables were selected to measure the behavioral attitudes, subjective norms and perceived behavioral controls of a migrant consistent with the Theory of Planned Behavior. The behavioral attitudes of a migrant largely relate to their sense of security (Smit et al. 2000; Smit and Wandel 2006; Brooks 2003; Handmer et al. 1999; Kelly and Adger 2000; IPCC 2007; Jones et al. 2010; de Sherbinin et al 2008; Perch-Nielsen 2004), thus the United Nation's Development Programme (UNDP) Broad Spectrum of Human Security Indicators (United Nations Development Programme 1994) was used to guide related parameter selection. Seven broad security categories are defined, including economic (wages, cost of living), food (caloric intake), health (infant mortality), environment (access to sanitation), personal (violent crime incidence, deaths due to conflict), community (migrant stocks in other countries), and political (corruption, disaster response). Perceived behavioral controls are related to an migrant's adaptive capacity, a convenient indicator being the wealth that a potential migrant holds, which can be broadly organized by financial capital (e.g., savings, credit, remittances), human capital (e.g., education, good health), social capital (e.g., interpersonal networks), physical capital (e.g., improved roads, communication), and natural capital (e.g., access to natural resources) (Brooks 2003; Handmer et al. 1999; Kelly and Adger 2000; IPCC 2007; Jones et al. 2010; de Sherbinin et al 2008; Perch-Nielsen 2004). Subjective norms capture the peer influence on the migration decision, which is treated here by the influence past migration behavior (lagged migration) has on the potential migrant's decision; that is, the inertia that past behavior has on current decisions.

Variables for the analysis were selected from the list of human security and capital stock measures, with particular attention paid to covering this broad range of factors important to the migration decision. Specific variable choices were greatly tempered by the completeness of the available data sets. Human security and capital stock data were taken largely from the World Bank's World Development Indicators database (World Band 2013). These data are available for 214 countries on an annual basis from 1960 to present, although the completeness of the data differs significantly between variable, country and year. Additional data sources include the International Disaster Database (Centre for Research on the Epidemiology of Disasters 2014), Polity IV Project (2014), and Battle Deaths Dataset (PRIO 2009). All data were normalized by population and referenced to the initial year of analysis, 1960. Once combined with the migrant stock data, only 166 countries were included in the model due to differences in the listing of countries across the various databases and general availability of data.

2.3.2. Economic Forecasts of International Migration

Econometric models remain a basic tool not only to predict migration, but also to verify certain economic theories based on empirical data. Most empirical research to date has focused on European circumstances. Our research extends the empirical literature on migration in the following ways: First, we augment the model of Hatton (1995) to include the costs of migration explicitly in our empirical model, such as the impact of disasters and battle deaths. Second, we take uncertainty in the model by considering home and destination country savings rate (Faini and Venturini, 1995). This implies that the sign of the home adjusted savings can be negative. Third, we include the variable polity levels of the home and the destination countries to account for non-economic factors in explaining outmigration patterns.

Several models have been proposed in the literature on the Theory of Planned Behavior in integrating micro-level and macro-level data of migration flows (Bijak, 2006). For example, the "Markov Chain model can be considered an appropriate representation of the structure of the behavioral process of repeat migrants" (Kupizewski, 2002). The problem with this model is that the homogeneity of populations under study and the stationarity of the stochastic process are artificial in the real-world model of migration determinants. Also, this kind of modeling framework requires detailed empirical data, which is usually not available. Weidlich and Haag (1988) developed a model that linked micro-level migratory decisions of individuals with their macro-level outcomes for inter-regional population flows. In their formulation a master equation method in statistical physics is used that relates the first order probability distributions of population over a vector of *N* states. These probability distributions are related to the differences between a single move of one person from the i^{th} state to the j^{th} state and the differences of the aggregate population from the i^{th} state to the j^{th} state. The transition rates were derived using a regression model that relates several socio-economic and distance variables on the migration process. Despite the potential usefulness of the model, the complexity of the model rendered it difficult to implement in empirical applications. Given the aggregate nature of the migration data in the present study, we propose a partial adjustment model that relates outmigration flows to its various socio-economic and political determinants. The method has its theoretical foundation in Hatton and Williamson (2001) and empirically formulated and implemented by Sinn (2001). As the objective of this method is to determine the elasticity of outmigration flows to its various determinants and then to validate the model using the state-space approach, we consider this method to be appropriate given the phenomena and the data structure. Thus, we first conduct a regression analysis of the determinants of outmigration flows, and then validate the model results with the state-space method.

A log-log form of the partial adjustment model was adopted (see Equation 1). This decision aided in interpreting the coefficients of the independent variables; that is, the coefficients in the log-log form of the model can be interpreted as elasticities. Additionally, the log-log form of the model makes the relationship between the dependent and independent variables linear, producing a more Gaussian series.

$$Ln(M_t) = \gamma[\beta_0 + \beta_1 Ln(GDP_t) + \beta_2 Ln(C_t) + \beta_3 Ln(D_t) + \beta_4 Ln(BD_t) + \beta_5 Ln(MO_t) + \beta_6 Ln(P_t) + \beta_7 Ln(T_t) + \beta_8 Ln(S_t) + \left(\frac{1}{\gamma} - 1 + \beta_9\right) Ln(M_{t-1}) + \varepsilon_t$$

(1)

Where M_t = (Migration Stock per capita) / (Migration Stock per capita in 1960)
GDP_t = (GDP per capita) / (GDP per capita in 1960)
C_t = (Cereal Production per capita) / (Cereal Production per capita in 1960)
D_t = (Disaster Impact per capita) / (Disaster Impact per capita in 1960)
BD_t = (Battle Deaths per capita) / (Battle Deaths per capita in 1960)
MO_t = (Infant Mortality per 1000) / (Infant Mortality per 1000 in 1960)
P_t = (Polity values in period X) / (Polity values in 1960)
T_t = (Miles of Telephone Lines per capita/miles of Telephone Lines per capita 1960)
S_t = (Per Capita Adjusted Savings / Per Capita Adjusted Savings 1960)

γ satisfies the long-run equilibrium relationship: $Ln(M_t) = Ln(M_{t-1}) + \gamma\ (Ln(M_t^*) —$ Ln(M_{t-1})); M_t^* denotes the equilibrium trajectory of foreign population stocks under study, where

$Ln(M_t^*) = \beta_0 + \beta_1 Ln(GDP_t) + \beta_3 Ln(C_t) + \beta_4 Ln(D_t) + \beta_5 Ln(BD_t) + \beta_6 Ln(MO_t) + \beta_7 Ln(P_t) + \beta_8 Ln(T_t) + \beta_9 Ln(S_t) + \beta_{10}\ Ln(M_{t-1})$
(2)

The value of γ was first obtained from the long-run equilibrium relationship. Then, equation (1) was solved to obtain the value of β_9. Substituting this back in equation (1) the adjustment coefficient$\left(\frac{1}{\gamma} - 1 + \beta_9\right)$ was determined.

Discussion of the Estimators
The two-dimensional nature of the panel data allows us to exploit both the variation between countries and time periods in the data for the estimation of the parameters of the migration function. The use of different estimators allows us to provide an answer to the methodological question on the extent of variation of the estimated coefficients across different estimation procedures.

We use the following estimators in our estimation procedure:

$LnM_{it} = \beta_0 + Ln(X'_{it})\beta_{it} + \omega_i + \lambda_t + \varepsilon_{it}$ (3)

where, M_{it} is the dependent variable, X_{it} is a *k*- vector of regressors, and ε_{it} are the error terms for each *i =1, 2,….M*, and *t* = 1, 2, …..T*,* where the subscript *i* denotes individual countries and *t* the dated periods. The β_0 parameter denotes the overall constant in the model, while the ω_i and λ_t represent the cross-section and period specific effects (random or fixed). Identification requires that the β coefficients have restrictions placed upon them. They may be divided into sets of common regressors (across cross-section and periods), cross-section specific regressors, and period-specific regressors. The dataset is unbalanced.

Four cases were considered:
1. Common β's: In this case equation (3) simplifies to

$$Ln\ M_{it} = \beta_0 + Ln(X'_{it})\beta + \omega_i + \lambda_t + \varepsilon_{it} \quad (4)$$

There are a total of *k* coefficients in β, each corresponding to an element of X.

2. Cross-Section Specific: In this case, equation (3) becomes

$$Ln\, M_{it} = \beta_0 + Ln(X'_{it})\beta_i + \omega_i + \lambda_t + \varepsilon_{it} \quad (5)$$

   There are a total of *MK* slope coefficients
3. Panel Generalized Least Squares (GLS) with common coefficients, and
4. Panel GLS with cross-sectional fixed effects.

The first specification (the pooled OLS estimator) uses both sources of variation (time and cross-sectional) in the data, although this estimator is not efficient. This specification ignores the individual heterogeneity of different countries, such as polity levels, that may have an important impact on migration flows.

The presence of cross-section and period specific terms ω_i and λ_t was handled using fixed and random effect methods. Models were specified in these ways individually and in both dimensions. Specifically, equation (1) was specified as a fixed effect in the cross-section dimension and a random effect in the period dimension. However, mixed effect specification was not pursued (for example, fixed effect in the cross-section and a random effect in the period dimension) because our dataset was unbalanced.

The random effects specification assumed that the corresponding effects ω_i and λ_t were realizations of independent random variables with mean zero and finite variance. Most importantly, the random effects specification assumed that the effect was uncorrelated with the residual term ε_{it}. Because the random effect estimation in periods generated large standard errors of the estimated coefficients, these estimates were not reported.

The model was also estimated using GLS to account for patterns of correlation between the residuals as given by cases 3 and 4. The variance structure considered was cross-section specific heteroskedasticity. This specification was chosen since residuals displayed non-normality while running least squares estimation. GLS corrected this residual variance for each cross-sectional unit. The GLS estimators are also superior in terms of efficiency when compared to other traditional types of estimators. This is achieved by the optimal weighting attached to the within and between variations in the data (Swamy and Arora, 1972). Although, both specifications III and IV are asymptotically equivalent, the estimated coefficients are likely to differ in both specifications.

In addition to the GLS efforts were made to test the GMM estimator developed by Arellano and Bover (1995) by applying the first difference transformation of the dependent and independent variables. Given that we had very few time periods, the lagged dependent variable as instruments did not produce results that were superior to the rest of the estimators used in this study.

2.3.3. Bilateral Model of Immigration

To complement emigration modeling, similar efforts were made to develop a model of bilateral immigration. That is, a model that would estimate where the emigrants (from the emigration model) would migrate. A similar approach was taken to the analysis; however, results were not what we had hoped. Details on the analysis and results can be found in Appendix B.

2.4. Results

Regression analysis results for the international emigration model, based on decadal data from 1960-2010, are given below. Specifically, results for a log-log model in which emigrant stocks for 166 countries around the world were regressed against the independent variables GDP, Cereal Production, Disaster Impact, Battle Deaths, Infant Mortality, Polity, Telephone Lines, and Adjusted Savings are given (see Appendix A for plots showing how the log transformation achieved Gaussian series). Additional analyses limited to developing and Asian nations are given for comparison. In efforts to validate these results a state-space analysis was conducted for the full 166 nation case. This involved deriving the fitted values of emigration flows for the 5-year incremental periods between the decadal data used in the regression analysis. Our results indicate that the state space approach gives identical results to the GLS specification with fixed effects. More importantly, our results confirm with the existing literature (Gu et al. 2014)] that the state space approach offers a computationally more efficient alternative to fit a generalized least squares regression model with a number of countries on each independent variable in a panel setting.

2.4.1. Regression Analysis

Regression analyses were performed using the commercial statistics package EViews, Version 8.0. Analyses included 830 unique measures (166 countries by 5 unique decadal periods [1960 was used as the referent case]). Multicollinearity was found between the independent variables, Adjusted Savings, Telephones Lines and GDP (positive correlation) and negative correlation between Adjusted Savings, Telephones Lines and Infant Mortality. Thus, we could not include these three pair of independent variables simultaneously in the regression model as this would bias results and give high standard errors of the independent variables. Thus, the following variables were included in a pairwise manner in order to account for the multicollinearity. They are GDP-Adjusted Savings; GDP-Telephone Lines; Infant Mortality-Adjusted Savings & Infant Mortality-Telephone Lines.

Results across the four model cases and between the four paired analyses (due to multicollinearity) yielded relatively similar results. Results are given in Table 2.1, which includes each of the model cases for the GDP-Telephone Lines pair. Given the similarity, result tables for the other three pairs have been moved to the Appendix A. Noteworthy across all cases and pairs was the strong fit between the model and data as indicated by R^2 values that range from 0.764 to 0.793 (higher values generally were associated with the GLS model cases).

Table 2.1: Regression Results for all countries model for the GDP and Telephone Lines Combination — (Dependent Variable: Log of emigration flows for the period 1970- 2010)

Variables	Common Coefficient	Cross-section Specific Fixed Effects	Panel GLS	Panel GLS with Cross-Section Fixed Effects

	I	II	III	IV
Constant	0.166*** (0.038)	0.193*** (0.043)	0.157*** (0.038)	0.186*** (0.045)
Log of Battle Deaths	0.005 (0.009)	0.009 (0.008)	0.006 (0.009)	0.009 (0.009)
Log of Cereal Production	0.084*** (0.029)	0.093*** (0.028)	0.083*** (0.026)	0.09*** (0.02)
Log of Natural Disasters	-0.001 (0.007)	-0.0012 (0.0078)	-0.001 (0.006)	-0.002 (0.007)
Log of GDP Index	-0.081** (0.037)	-0.087** (0.037)	-0.081** (0.034)	-0.085** (0.034)
Log of Lagged Outmigration	1.059*** (0.018)	1.054*** (0.02)	1.053*** (0.019)	1.048*** (0.02)
Log of Polity Index	0.02 (0.05)	0.009 (0.05)	0.029 (0.046)	0.019 (0.046)
Log of Telephone lines per capita	-0.028* (0.015)	-0.039** (0.017)	-0.027* (0.014)	-0.036** (0.016)
Adjusted R^2	0.773	0.773	0.793	0.793
N	782	782	782	782
DW Statistic	1.946	1.953	1.96	1.965

Across the four model cases and four paired analysis, coefficients for Lagged Migration, GDP, Telephone Lines, Adjusted Savings, Infant Mortality and Cereal Production were significant at the 1-5% level in all but one case (Adjusted Savings for the Infant Mortality-Adjusted Savings pair). Specifically, the elasticity of emigration with respect to GDP was negative and significant lying between -0.137 and -0.1. This result may suggest that as income of the country of origin declines, citizens are more likely to emigrate abroad. The elasticity of emigration with respect to Telephone Lines was negative and significant lying between -0.039 and -0.027. This result may suggest that as the infrastructure decays the interest in emigration increases. In contrast, Lagged Migration had a positive and significant effect with values between 1.048 and 1.081. This result may suggest that as more citizens choose to emigrate, it is also likely that family members and friends are likely to follow. The elasticity of emigration with respect to Infant Mortality was between 0.064 to 0.088, depending on the model specification. This result may suggest that as the Infant Mortality rate of the country of origin increases, individual's perception of their health security declines which leads them to migrate abroad. The elasticity of emigration with respect to Cereal Production was positive and significant lying between 0.066 to 0.078. This result may suggest that an increase in cereal production may enable emigration by increasing resources for certain sectors of the population, while also increasing the motivation to migrate by other sectors of the population whose access to arable land is impacted by commercialization of cereal production. Finally, the elasticity of emigration with respect to Adjusted Savings was positive and significant lying between 0.0005 and 0.015 suggesting that as an individual's adaptive capacity as measured by Adjusted Savings improves their tendency to emigrate increases.

The goodness of fit between the emigration model and the measured emigration data was explored graphically. Several nations were selected that represent a range in the visual degree of fit (Figure 2.1).

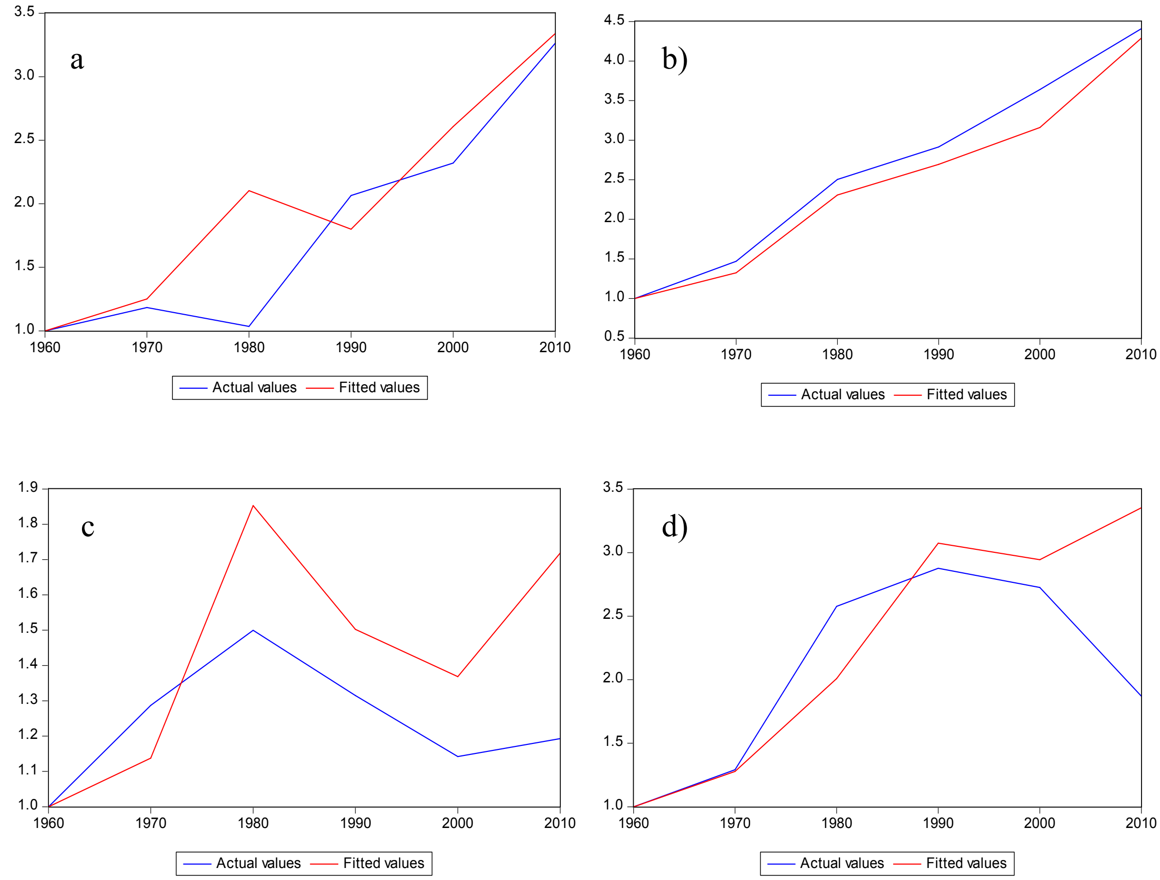


Figure 2.1. Actual verses fitted values of emigration flows (1960-2010) for a) Slovenia, b) Columbia, c) Tunisia, and d) Panama.

The emigration model was further explored by focusing the regression analysis on specific nations; specifically, developing countries (citation) and countries comprising the continent of Asia. The goal was simply to identify potential differences across these countries. Results for developing nations are given in Table 2.2, again limited to the paired analysis for GDP and Telephone Lines (consistent with Table 2.1). Results for the other model cases and paired analyses are included in the Appendix A. Goodness of fit is slightly improved over the all nations case (R^2 between 0.772-0.81). Similarities with the all nations analysis included the elasticity of emigration was negative and significant for Telephone Lines and positive and significant for Lagged Migration and Cereal Production (elasticities of Cereal Production were slightly higher 0.091-0.124). Alternatively, significant differences were noted in the fact that neither GDP, Infant Mortality nor Adjusted Savings registered a statically significant connection with emigration. This could suggest that conditions tend to be challenging in developing countries and as such, changes in these factors have little influence on the migration decision.

The emigration model for the Asian nations was seen to deviate significantly from that of the all nations' and developing nations' case (Table 2.3). Only Lagged Migration was found to be consistently correlated (ranging from 0.994 to 1.064) to the emigration data across all cases and pairings. In contrast, Infant Mortality showed a significant and positive trend, while unique to the Asian nations Battle Deaths was seen to be positive and significant but only for a limited set of models and pairings. These differences undoubtedly reflected the reduced number of countries included in the regression analysis but also reflected the unique characteristics of Asian countries. Consistent with the other regression analyses, the goodness of fit was strong as the R^2 ranged from 0.77 to 0.853.

2.4.2. Validation Using a State-Space Model

Additional efforts were made to validate the all nations emigration model. This was pursued by developing a state-space model to independently estimate emigration flows. The state-space model was selected for the validation studies because: 1) it integrates unobserved components called state variables with observable series in a single system, and 2) it uses a recursive algorithm called Kalman filtering to recursively update the state variables. The study used decadal level emigration flows and treated its five-year incremental values as missing observations. Specifically, using a state-space model, a time varying parameter model with the signal equation error term specified as an AR(1) process, the five-year predicted values of emigration flows were obtained. Details of the state-space model application are included in the Appendix to this paper.

Results were first compared by excluding the Lagged Migration flows as a state-space variable and looking at the influence of Infant Mortality, Battle Deaths, GDP and Adjusted Savings on current emigration flows. Then, Lagged Migration flows were included in the state-space modeling framework to show how model fit improves. These results are shown in Tables 2.4 and 2.5.

Table 2.4 indicates that all variables are highly significant, namely GDP, Adjusted Savings, Infant Mortality, and Battle Deaths. However, excluding the Lagged Migration results in poor model fit as given by the value of the log likelihood. This may suggest that including the Lagged Migration flows is desirable in each model specification.

Table 2.2: Regression Results for developing countries model with GDP and Telephone Lines Combinations — (Dependent Variable: Log of emigration flows for the period 1970-2010)

Variables	Common Coefficient I	Cross-section Specific Fixed Effects II	Panel GLS III	Panel GLS with Cross-Section Fixed Effects IV
Constant	0.179*** (0.042)	0.206*** (0.047)	0.174*** (0.041)	0.21*** (0.049)
Log of Battle Deaths	0.002 (0.012)	0.006 (0.012)	0.003 (0.011)	0.006 (0.011)
Log of Cereal Production	0.093** (0.047)	0.105** (0.046)	0.097** (0.048)	0.111** (0.049)
Log of Natural Disasters	-0.001 (0.007)	-0.002 (0.007)	-0.003 (0.007)	-0.003 (0.007)
Log of GDP Index	-0.028 (0.036)	-0.036 (0.038)	-0.035 (0.037)	-0.041 (0.038)
Log of Lagged Outmigration	1.042*** (0.023)	1.035*** (0.025)	1.032*** (0.022)	1.025*** (0.024)
Log of Polity Index	0.08 (0.062)	0.074 (0.063)	0.08 (0.06)	0.076 (0.06)
Log of Telephone lines per capita	-0.033** (0.015)	-0.045** (0.02)	-0.028* (0.015)	-0.042** (0.02)
Adjusted R^2	0.782	0.783	0.809	0.81
N	626	626	626	626
DW Statistic	1.948	1.96	1.956	1.97

Notes: ***, **, and * denotes significance at 1%, 5%, and 10% respectively.
The terms in the brackets denotes standard errors.

Table 2.3: Regression Results for the Asia model with GDP and Telephone Lines Combinations — (Dependent Variable: Log of emigration flows for the period 1970- 2010)

Variables	Common Coefficient I	Cross-Section Specific Fixed Effects II	Panel GLS III	Panel GLS with Cross-Section Fixed Effects IV
Constant	0.255** (0.119)	0.149 (0.11)	0.238** (0.105)	0.196* (0.104)
Log of Battle Deaths	0.006 (0.007)	0.013 (0.009)	0.013* (0.007)	0.018* (0.01)
Log of Cereal Production	0.014 (0.021)	0.021 (0.019)	0.025 (0.035)	0.026 (0.032)
Log of Natural Disasters	-0.007 (0.021)	-0.0009 (0.022)	-0.012 (0.075)	-0.007 (0.019)
Log of GDP Index	-0.011 (0.082)	-0.055 (0.092)	-0.012 (0.075)	-0.039 (0.082)
Log of Lagged Outmigration	1.018*** (0.05)	1.045*** (0.047)	0.994*** (0.045)	1.012*** (0.047)
Log of Polity Index	0.131 (0.198)	0.186 (0.212)	0.125 (0.181)	0.171 (0.2)
Log of Telephone lines per capita	-0.042 (0.027)	0.015 (0.032)	-0.026 (0.181)	0.009 (0.031)
Adjusted R^2	0.775	0.778	0.853	0.852
N	163	163	163	163
DW Statistic	1.73	1.79	1.791	1.84

Notes: ***, **, and * denotes significance at 1%, 5%, and 10% respectively.
The terms in the brackets denotes standard errors.

Table 2.4: State-space model estimation with GDP, Adjusted Savings, Infant Mortality and Battle Deaths as State Space Variables (Lagged Migration excluded)

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	1.247681	0.093715	13.31356	0.0000
C(7)	-1.250806	0.022739	-55.00696	0.0000
C(8)	0.041808	0.032055	1.304274	0.1921
	Final State	Root MSE	z-Statistic	Prob.
GDP	-0.402457	0.058684	-6.858032	0.0000
Infant Mortality	-0.283188	0.057820	-4.897718	0.0000
Battle Deaths	0.129276	0.044269	2.920252	0.0035
Adjusted Savings	0.108535	0.024660	4.401233	0.0000
SV6	0.020022	0.535048	0.037421	0.9701
Log likelihood	-803.3616	Akaike info criterion		1.669486
Parameters	3	Schwarz criterion		1.684620
Diffuse priors	5	Hannan-Quinn criter.		1.675247

Notes: SV6 is a state variable; C(1) , C(7) and C(8) are defined in sspace05 specification in supplementary materials file.

In Table 2.5, Infant Mortality, Lagged Migration and Cereal Production were included as state-space variables. GDP was excluded from this model. All variables were found to be highly significant suggesting that the state-space model did a good job in validating the panel GLS model as evidenced in terms of by a higher log likelihood value. In the GLS model with infant mortality, cereal production and lagged migration as independent variables (Table A.3 of Appendix A), all these variables were found to be highly significant in explaining emigration flows. The state space model also corroborates this finding using the Kalman filter (KF) algorithm. Calculated coefficients are not only highly significant but it is also easier and more efficient to implement this algorithm over the GLS model. Thus, the KF algorithm is more efficient than the algorithms specifically designed for univariate and multilevel regression modeling (Bauer 2003; Curran 2003). In addition, the model also predicted the intermediate values of the emigration flows for the periods 1965, 1975, 1985, 1995 and 2005.

Algorithms of the state-space model with errors represented as state variable with AR(1) process and the actual and smoothed forecasted values of outmigration flows for countries using GDP, Cereal Production, Lagged Migration as state-space variables for the Year 1980 are included in the Appendix A.

Results indicate that the state-space approach is essentially identical to the GLS with fixed effects specification. This is consistent with the existing literature (Gu et al. 2014; Bauer 2003; Curran 2003). More importantly, the state space approach offers a better way to fit the fixed effects GLS model with a large number of countries on each independent variable in a panel data setting.

Table 2.5: Effect of Infant Mortality and Lagged Migration Flows on current emigration flows

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	-3.039798	0.025705	-118.2576	0.0000
C(2)	0.997467	0.001734	575.2135	0.0000
	Final State	Root MSE	z-Statistic	Prob.
Cereal Production	0.085064	0.033894	2.509691	0.0121
Infant Mortality	0.146010	0.044215	3.302295	0.0010
Lagged Migration	1.827321	0.435262	4.198212	0.0000
SV4	3.125509	0.310216	10.07526	0.0000
Log likelihood	-260.6417	Akaike info criterion		0.544335
Parameters	2	Schwarz criterion		0.554433
Diffuse priors	4	Hannan-Quinn criter.		0.548180

Notes: SV4 is a state variable; C(1) and C(2) are defined in sspace07 specification in supplementary materials file.

2.5. Discussion and Conclusions

A dynamic model of international human emigration was developed. The model was structured according to the Theory of Planned Behavior. This national-level model integrated dynamics of human security, adaptive capacity, and social interaction that inform the migration decision. This emigration model was designed to operate within the broader context of integrated assessment modeling of global climate change.

The model was established through regression analysis utilizing historical migration data for 166 countries. Specifically, dependent variable time series data, 1960-2010, were extracted from the bilateral migrant stock data published by the World Bank, while associated independent variable time series data were largely taken from the World Bank's World Development Indicators database. Results for the all nations emigration model yielded a surprisingly strong fit to historical data (R^2 on the order of 0.8). Coefficients for Lagged Migration, GDP, Telephone Lines, Adjusted Savings, Infant Mortality and Cereal Production were significant at the 1-5% level. Correlation between GDP, Adjusted Savings, Infant Mortality and Telephone Lines was noted.

The positive coefficients for Infant Mortality, which suggests the importance of access to health care on emigration and the negative coefficients for Telephone Lines, which suggest decaying infrastructure encouraged migration, both reinforce the understanding that people migrate in hopes of improving their quality of life (Ravenstein 1889; Massey et al. 1993; Black et al. 2011). The positive coefficients for Lagged Migration, which indicate the influence of prior decisions by friends and family, are consistent with multiple empirical studies that demonstrate the influence of networks on emigration (Massey et al. 1993; Davis et al. 2013; Massey 1988; Clark et al 2007).Positive coefficients for Adjusted Savings suggest the importance of financial capital on international emigration consistent with multiple accounts of the cost of international migration (Black et al. 2011; Laczko and Aghazarm 2009; Piguet et al. 2011).The negative coefficients for GDP indicate declining economic conditions promoted migration. This finding is consistent with findings by Reuveny and Moore (Reuveny and Moore 2009)] but inconsistent with Afifi and Warner (2008) who found that GDP per capita had a positive impact on migration.

The positive coefficients for Cereal Production suggest that increases in cereal production may enable and motivate emigration by increasing resources for certain sectors of the population, while decreasing opportunities for others. This finding is inconsistent with Feng and others (2010) and Halliday (2006) who show that adverse agricultural conditions increased international migration. However, it appears to be consistent with the finding from Gray (2008) that international migration is most likely from land-rich households as well as the findings by Findley (1994), Henry and others (2004), and van der Geest (2008) that community members wait for improved agricultural and economic conditions before migrating oversees. This finding is also consistent with Reuveny and Moore (2009), who found that an increase in cropland had a positive effect on out-migration. The increase in cropland and corresponding decrease in arable land, they suggest, indicates a lack of agricultural opportunities for non-land holders that may motivate migration decisions.

Battle Deaths and Disaster Impacts were both absent from the list of significant variables contributing to emigration. These variables were originally adopted to explore the potential impacts of conflict and climate change, respectively, on the emigration decision. The reason for this absence may lie, in part, in the 10-year frequency of the migration stock data; that is, the effects of conflict and climate are occurring at a higher frequency than can be detected with the 10-year data. The absence of battle deaths from the list of significant variables may also reflect the role of conflict in both motivating migration and preventing people from leaving (Black et al. 2011; Lubkemann

2009). Reuveny and Moore (2008) found that while civil war had a positive effect on out-migration, war, which may increase incentives to stay and defend the homeland and increase the difficulty of moving, had a negative effect (Reuveny and Moore 2008). The absence of disaster impacts from the list of significant variables is inconsistent with Naude (2008) and Reuveny and Moore 2008), but consistent with multiple studies that show disasters typically trigger short-term internal migrations that would not register in this decadal assessment of international migration (Internal Displacement Monitoring Center 2011; Piguet et al. 2011; Halliday 2006; Gray 2008; Paul 2005).

Regression analyses were conducted for subsets of nations, specifically developing nations and Asian nations. Each subset of nations yielded similarly good fit between the model and historical emigration trends. Also common to all countries was the significant and positive influence of Lagged Migration on the emigration decision. In contrast, each community of nations exhibited differences in terms of the resulting coefficients. For developing nations, GDP, Infant Mortality and Adjusted Savings were no longer found to contribute significantly to the emigration decision. As the community became more focused (Asian nations) differences with the all nations model became more evident as only the coefficients for Lagged Migration were significant. While these differences likely reflect statistical differences due to the sample size used in the different regression analyses, these results are also likely to reflect differences in the dynamics influencing the emigration decision across these three groupings of countries.

The expansion of our analysis of the panel GLS model of the log-log specification using a state-space model is a major innovation in the international development literature. As mismatch in frequency remains a problem for econometricians for quite some time, this paper contributes to the existing literature on international development by modeling the outmigration flows in a state-space modeling framework. As temporal aggregation may result in information loss, mixed frequency model may be a good alternative to the existing modeling paradigm.

Since this field is relatively new and evolving, the collection of literature is also fairly limited. While mixed frequency models have been used to some extent in financial engineering, macroeconomic forecasting and other related disciplines, to the best of our knowledge this is the first study that accommodates state space modeling framework to an international development problem.

2.6. References

Ravenstein, E.G., 1889. The Laws of Migration, *Journal of the Royal Statistical Society*, Vol. 52, pp 241-310.

Internal Displacement Monitoring Center and the Norwegian Refugee Council, 2011. *Internal Displacement: Global Overview of Trends and Developments in 2010.*

Salehyan, I., and Gleditsch, K.S. , 2006. Refugees and the Spread of Civil War, *International Organization,* Vol. 60, pp 335–366.

Massey, D.S. Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., and Taylor, J.E., 1993. Theories of Migration: A Review and Appraisal, *Populating and Development Review*, Vol. 19, No. 3, pp 431-466.

Ghimire, D.J. and Mohai, P.L., 2005. Environmentalism and Contraceptive Use: Howe People in Less Developed Settings Approach Environmental Issues, *Population and Environment*, Vol. 27, No. 1, pp 29-61.

Davis, K.F., D'Odorico, P., Laio, F., Ridolfi, L., 2013..Global Spatio-Temporal Patterns in Human Migration: A Complex Network Perspective, *PLoS ONE*, Vol. 8, No. 1.

Grote, U., Engel, S., Schraven, B., 2006. *Migration due to the Tsunami in Sri Lanka –Analyzing Vulnerability and Migration at the Household Level*. ZEF-Discussion Paper on Development Policy No. 106, Centre for Development Research (ZEF), Bonn, April 2006.

El-Hinnawi, E., 1985. *Environmental Refugees*. United Nations Environment Programme, Nairobi.

Myers, N., 2002. Environmental Refugees: A Growing Phenomenon of the 21st Century, *Philosophical Transactions of the Royal Society B*. Vol. 357, pp 609-613.

Brown, O., 2008. *Migration and Climate Change, International Organization for Migration (IOM): Research Series No. 31*, IOM Geneva.

Christian Aid, 2007. *Human Tide: The real migration crisis,* Christian Aid Report May 2007, London.

CNA, 2009. National Security and the Threat of Climate Change.

Defence Science Board, 2011. Trends and Implication of Climate Change for National and International Security.

National Intelligence Council, 2012. Global Trends 2030: Alternative Worlds. NIC 2012-001.

United Nations Framework Convention on Climate Change, 2013. Nairobi work programme on impacts, vulnerability and adaptation to climate change, Decision 17/CP.19.

Lee, E., 1966. A Theory of Migration, *Demography*, Vol. 3, No. 1.

McLeman, R., and Smit, B., 2006. Migration as an Adaptation to Climate Change. *Climatic Change*, Vol. 76, No.1-2, pp 31-53.

Black, R.B., Adger, N., Arnell, N., Dercon, S., Geddes, A., Thomas, D., 2011. *Foresight: Migration and Global Environmental Change, Final Project Report*, The Government Office for Science, London.

Gilbert, G., and McLeman, R., 2010. Household access to capital and its effects on drought adaptation and migration: A case study of rural Alberta in the 1930s. *Population and Environment*, Vol.32, No.1, 3-26.

Faist, T., 2000. *The volume and dynamics of international migration and transnational social spaces*. Oxford: Clarendon Press.

Afifi, T. and Warner, K., 2008. *The Impact of Environmental Degradation on Migration Flows Across Countries, United Nations Working Paper No. 5/2008*, Bonn.

Reuveny, R., and Moore, 2009. W.H., Does Environmental Degradation Influence Migration? Emigration to Developed Countries in the Late 1980s and 1990s, *Social Science Quarterly*, Vol. 90, No. 3, pp 461-479.

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

Walsh, S. J., Malanson, G. P., Entwisle, B., Rindfuss, R. R., Mucha, P. J., Heumann, B. W., McDaniel, P. M., Frizzelle, B. G., Verdery, A. M., Williams, N. E., Yao, X., and Ding, D., 2013. Design of an agent-based model to examine population-environment interactions in Nang Rong District, Thailand. Applied Geography 39:183-198.

Zhang, Q. and Jager, W., 2011. Agent based modeling of population dynamics in municipalities: Migration in the Derbyshire and Nottinghamshire cases in the UL, Report for the EU Prima project, WP3, December 2011.

Hassani-Mahmooei, B., and Parris, B. W., 2012. Climate change and internal migration patterns in Bangladesh: an agent-based model. Environment and Development Economics 17:763-780.

Parson, E.A. and Fisher-Vanden, K.,1997. Integrated Assessment Models of Global Climate Change, *Annu. Rev. Energy Environ*. 1997. 22:589–628.

Ajzen, I., 1991. The theory of planned behavior. Organizational Behavior and Human, *Decision Processes*, Vol. 50, pp 179–211.

Kniveton, D., Smith, C. and Wood, S., 2011. Agent-based model simulations of future changes in migration flows for Burkina Faso. Global Environmental Change 21:S34-S40.

Kniveton, D.R., Smith, C.D., Black, R., 2012. Emerging Migration Flows in a Changing Climate in Dryland Africa, Nature Climate Change Letters, published online March 11, 2012.

De Jong, G., 1999. Choice processes in migration behaviour. IN PANDIT, K. & WITHERS, S. D. (Eds.) Migration and Restructuring in the U.S. New York, Rowman and Littlefield Publishers, Inc.

De Jong, G., 2000. Expectations, gender, and norms in migration decision-making. *Population Studies,* 54**,** 307-319.World Bank. 2013. World Development Indicators 2013. Washington, DC: World Bank.

Smit, B., Burton, I. Klein, R.J.T., and Wandel, J., 2000.An Anatomy of Adaptation to Climate Change and Variability, *Climatic Change*, Vol. 45, pp 223-251.

Smit, B. and Wandel, J., 2006. Adaptation, Adaptive Capacity and Vulnerability, *Global Environmental Change,* Vol. 16, pp 282-292.

World Bank and Ç. Özden, C. Parsons, M. Schiff and T. L. Walmsley, 2011. Where on Earth is Everybody? The Evolution of Global Bilateral Migration, 1960-2000. World Bank Economic Review 25(1):12-56.

Brooks, N., 2003. *Vulnerability, Risk and Adaptation: A conceptual Framework. Working Paper 38*, Tyndall Centre for Climate Change Research, University of East Anglia, Norwich.

Handmer, J.W., Dovers, S., Downing, T.E., 1999. Societal Vulnerability to Climate Change and Variability, *Migration and Adaptation Strategies for Global Change*, Vol. 4, pp 267-281.

Kelly, P.M., and Adger, W.N. 2000. Theory and Practice in Assessing Vulnerability to Climate Change and Facilitating Adaptation, *Climate Change*, Vol. 47, pp 325-352.

IPCC, 2007. *Climate Change Impacts, Adaptation and Vulnerability*, Geneva: IPCC.

Jones, L., Ludi, E., and Levine, S., 2010. *Towards a Characterisation of Adaptive Capacity: A Framework for Analysing Adaptive Capacity at the Local Level*, Overseas Development Institute: Background Note, December 2010.

de Sherbinin, A., VanWey, L. K., McSweeney, K., Aggarwal, R. Barbieri, A., Henry, S., Hunter, L. M., Twine, W. and Walker, R., 2008. Rural household demographics, livelihoods and the environment. Global Environmental Change-Human and Policy Dimensions 18:38-53.

Perch-Nielsen, S.L., 2004. *Understanding the effect of climate change on human migration: The contribution of mathematical and conceptual models, diploma thesis*. Swiss Federal Institute of Technology, Zurich.

United Nations Development Programme, 1994. *Human Development Report*, Oxford University Press, New York.

World Bank. 2013. World Development Indicators 2013. Washington, DC: World Bank.

Centre for Research on the Epidemiology of Disasters, 2014. The International Disaster Database, available at: http://www.emdat.be/

Polity IV Project, 2014. Political Regime Characteristics and Transitions, 1800-2013, available at: http://www.systemicpeace.org/polity/polity4.htm

PRIO, 2009. The Battle Deaths Dataset version 3.0, available at: http://www.prio.org/Data/Armed-Conflict/Battle-Deaths/The-Battle-Deaths-Dataset-version-30/

Sinn, H.-W., Flaig, G., Werding, Munz, M. S., D¨ull, N. and Hoffmann, H., 2001. EU-Erweiterung und Arbeitskr¨ aftemigration. Wege zu einer schrittweisen Ann¨aherung der Arbeitsm¨arkte. M¨unchen: Ifo-Institut f¨ur Wirtschaftsforschung.

Gu, F., Preacher, K.J. Wu, W. and Yung, Y-F., 2014. A Computationally Efficient State Space Approach to Estimating Multilevel Regression Models and Multilevel Confirmatory Factor Models, *Multivariate Behavioral Research*, 49, pp. 119-129.

Bauer, D.J., 2003. Estimating multilevel linear models as structural equation models, *Journal of Educational and Behavioral Statistics*, 28, pp.135-167.

Curran, P.J., 2003. Have multilevel models been structural equation models all along? Multivariate Behavioral Research, 38, pp. 529-569.

Massey, D., 1988. Economic development and international migration in comparative perspective, *Population and Development Review,* Vol. 14, No. 3, pp. 383-413.

Clark, X., Hatton, T.J., 2007. Williamson, J.G., Explaining US immigration 1971-1998, *Review of Economic Statistics*, Vol. 89, pp. 359-373.

Laczko, F. and Aghazarm, C. (eds.), 2009. *Migration, Environment and Climate Change: Assessing the Evidence*, Geneva: IOM.

Piguet, E., Pecaoud, A., and de Guchteneire, P., 2011. Migration and climate change: an overview, *Refugee Survey Quarterly*, pp. 1-23.

Feng, S., Krueger, A.B., Oppenheimer, M., 2010. Linkages among climate change, crop yields and Mexico-US cross-border migration, *Publication of the National Academies of Science (PNAS),* Vol. 107, No. 32, pp. 14257-14262.

Halliday, T. 2006. Migration, risk, and liquidity constraints in El Salvador, *Economic Development and Cultural Change*, Vol. 54, No. 4, pp. 893-925.

Gray, C., 2008. Environment, land and rural out-migration in the southern Ecuadorian Andes, *World Development*, Vol. 37, No. 2, pp. 457-468.

Findley, S. E., 1994. Does drought increase migration? A study of migration from rural Mali during the 1983-1985 drought, *International Migration Review*, Vol. 28, No. 3, pp. 539-553.

Henry, S., Schoumaker, B., and Beauchemin, C., 2004. The impact of rainfall on the first out-migration: a multi-level event-history analysis in Burkina Faso, *Population and Environment*, Vol. 25, No. 5, pp. 423-460.

van der Geest, K. 2008. North-South migration in Ghana: what role for the environment? *International Migration*, Vol. 49, No. 51, pp. 69-94.

Lubkemann, S.C., 2009. Involuntary mobility: on a theoretical invisibility in forced migration studies, *Journal of Refugee Studies*, Vol. 21, No. 4, pp. 454-475.

Naude, W., 2008. Natural disasters and international migration from sub-Saharan Africa, *Migration Letters*, Vol. 6, No. 2, pp. 165-176.

Paul, B.K., 2005. Evidence against disaster-induced migration: the 2004 tornado in north-central Bangladesh, *Disasters*, Vo. 29, No. 4, pp. 370-385.

Kupiszewski, M. 2002. How trustworthy are forecasts of international migration between Poland and the European Union? *Journal of Ethnic and Migration Studies*, 28(4), pp. 627-645.

Weidlich, F., and Haag, G. (eds.), 1988. Interregional migration: dynamic theory and comparative analysis, Springer, Berlin-Heidelberg.

Swamy, P.A.V.B, and Arora, S.S., 1972. The exact finite sample properties of the estimators of coefficients in the error component regression models, *Econometrica*, 40, pp.261-275.

Hatton, T.J. 1995. A model of UK migration 1870-1913, *Review of Economics and Statistics*, 77(3), pp. 407-415.

Bijak, J. 2006. Forecasting International Migration: Selected Theories, Models, and Methods, *Central European Forum for Migration Research* Working Paper No. 4/2006.

Harvey, A.C. 1989. *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge, Cambridge University Press.

Tsimikas, J.V. and Ledolter, J. 1997. Mixed Model Representation of State Space Models: New Smoothing Results and Their Application to REML Estimation, *Statistica Sinica*, 7, pp. 973-991.

Faini, R. and Venturini, A. 1995. *Migration and Growth: The Experience of Southern Europe,* CEPR Discussion Paper No. 964.

Arellano, M. and Bover, O. 1995. Another look at the Instrumental variable estimation of error-components models, *Journal of Econometrics*, 68, pp. 29-52.

3. LONG-TERM MIGRATION DYNAMICS

3.1. Basic Concepts

This LDRD originated from a concern that climate change would stress local areas and induce migration. The local loss of income combined with the effectiveness of or lack of constructive government response can hypothetically lead to violence that can spillover across a region.[1] Africa and the Middle East are areas where climate change could most likely lead to societal stresses. Mali became a centroid of the analysis because it was not burdened with predisposed notions of associations. (For background on the Mali conflict, see Francis 2013).

Chapter 2 explored those variables that affect long-term migration by considering a model of migration push. The concept of "push migration" is the consideration of those conditions or characteristics that affect the rate at which individuals emigrate from a country. These same variables are then also relevant to a push-pull migration. The concept of "pull migration" is the consideration of those conditions or characteristics that affect the rate at which individuals immigrate to a country. A push-pull model of migration attempts to capture both sets of characteristics simultaneously in an effort to understand the migration flows among a collection of counties. Although conventional methods of analysis can consider push-pull dynamics, it is difficult to obtain adequate correspondence between model and data, as noted in the appendices discussed in Chapter 2. This chapter presents an alternative approach to address push-pull migration dynamics. Push-pull migration considers the relative nature of choices for would-be migrants. It compares the utility of the existing location to the utility of the alternative locations. More of this logic is provided in the appendix on the model design.

The long-term aspect of migration, as used here, designates an indeterminate or permanent intention to move to and stay within another location. Chapter 4 of this report considers short-term, temporary migration. Short-term migration is dominantly a push process driven largely by presumed temporary conditions within the home locale, such as military activity, local violence, or natural disaster. Chapter 5, before the discussion of the complete modeling framework, presents preliminary, but promising, results on spillover violence.

3.2. Analysis Approach

Sandia has recently developed a robust approach for simulating the behaviors of individuals and groups called Behavioral Influence Assessment (BIA) (Barnard et al, 2014, and Backus et al, 2010). BIA modelling considers the reality that people are sensitive to certain types of information which they filter and form into patterns for making decisions. Many times the patterns are used to compare current conditions to remembered or expected conditions. Discordance between the perceived current conditions and the expected conditions can be a major driver of conscious or subconscious decisions, further amplified by emotively-charged responses. Long-term migration is a decision that should be amenable to simulation using the

[1] http://www.aljazeera.com/indepth/features/2015/04/climate-change-food-shortages-conflict-mali-150426105617725.html

BIA construct. Spillover violence is also the result of decisions that could be adequately described through a BIA-type simulation.
The mathematical approach for statistically estimating the parameters of such behaviors fall under Qualitative Choice Theory (Backus and Glass 2006, McFadden 1974, 1982, 1986). We use this approach to characterize the functional form and statistical analysis demonstrated in the next three chapters.

3.3. Data

The push-pull migration analysis separates Mali into Urban and Rural areas. It then includes the Neighboring countries in aggregate (CIV Ivory Coast, GAB Gabon, GHA Ghana, GMB Gambia, GNB Guinea-Bissau, MRT Mauritania, NER Niger, NGA Nigeria, and SEN Senegal). The remaining destinations are US and Rest of World (ROW). An aggregate of Western European is used as a relevant proxy for ROW conditions. The modeling calculates the choice to migrate, with the result being the recorded share of a country-of-origin's migrants in a given host country. The historical bilateral migration data from 1960 to 2013 data comes from the World Bank[2]. The referent population is the total of all Malians living world-wide, not just those living in Mali.

Based on the work of the previous chapter, the information considered as part of the decision calculus for the population, in the push-pull analysis, includes:

Per Capita Income - from the World Bank (WB) World Development Indicators (WDI) dataset[3] in real US dollars.

Food availability - calculated as the WB WDI Food Product Index divided by in-country population

Government Effectiveness— The Control of Corruption indicator from the World Bank Governance Indicators (WGI) dataset[4] as a proxy for government effectiveness within the country (rescaled to a range of 0 to 100). The lower it is, the more corrupt the population perceives the government.

Government Infrastructure: The Regulatory Quality indicator from the World Bank Governance Indicators (WGI) dataset as a (inverse) proxy for government infrastructure (rescaled to a range of 0 to 100). The higher it is, the better the government is able to promote private sector development.

Disease Multiplier – using the WB WDI under 5 years-of-age mortality rate.

Violence Expectation – the remembered level for Rule of Law from the WGI using an exponential filter of 5 years. The actual value was also tested, as will be discussed later. The lower the value, the greater the perceived violence.

Income Expectation– the remembered level income using an exponential filter of 5 years.

---

[2] http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTDECPROSPECTS/0,,contentMDK:22759429~pagePK:64165401~piPK:64165026~theSitePK:476883,00.html
[3] http://data.worldbank.org/data-catalog/world-development-indicators
[4] http://data.worldbank.org/data-catalog/worldwide-governance-indicators

Food Expectation– the remembered level food availability using an exponential filter of 5 years.

Labor Participation Rate – from WB WDI for males between 15 and 64. This value subtracted from 1.0 to act as a proxy for unemployment or conversely employment opportunity. Actual unemployment data is too sparse to use.

Population – the origin-population (Malians) divided by the host country population from WB WGI. This fraction is a proxy to capture the support group a migrant may have in a new host country. The rural versus urban split in Mali uses the WB WDI data.

Natural Disaster – fraction of population affected by natural disasters using the International Disaster Database from the Centre for Research on the Epidemiology of Disasters (CRED). This is a hook to include extreme climate change conditions. The actual simulation noted in a later chapter, captures climate impacts as detailed in the simulation model appendix. The expected value was also tested, as will be discussed later.

This set of dependent variables will be used to estimate the migration utility for each area. The index used to designate this set is noted as "k" in the subsequent discussion.

The simulation model includes remittances in income, and climate impacts within the economy via consequence to food production, water availability for other economic activities, and labor productivity. This approach imposes climate conditions on the choice decision through a direct causal chain rather than by correlation. It departs from the conventional approach for statistical analysis of climate impacts (Coniglio and Pesce 2015) that incorporates all climate and demographic conditions into a single equation.

The data covers the periods from 1960 to 2013. Missing data are interpolated where possible, but when missing data cannot be interpolated between missing entries, the data is set to an average historical, neutral, or zero value, as appropriate, for minimizing the biasing impact on parameter estimation. Standard imputation methods were not used because of the strong autocorrelation of the relevant variables over the time period.

The WGI indicators are from surveys and represent perceptions. The analysis uses them rather than any objective measures of violence or governance conditions because perceptions are what the decision maker, a human individual, uses to make the migrate choice.

3.4. Mathematical Construct

In a QCT construct, the indicated migration population (MP) is determined by:

$$MP_{i,l,g,v} = Exp(MU_{i,l,g,v}) / \sum_{k} Exp(MU_{k,l,g,v})$$

Where MU is the migration utility. The indexing is designed to cover region (i), labor classification (l), gender (g), and age (v). To demonstrate the applicability of the methods, this discussion simply focused on the country choice over the aggregate of population. The utility is restricted to only use information (or some filtered variant of it) to which the population actually has access. Discordance can be represented as perceived current information divided by the expected information. If the dependent variables are recognized by decision makers through relative, rather than absolute differences, then a logarithmic formulation is appropriate and the dissonance is the logarithm (ln) of the current value less the expected value.

The Migration Utility (using only "i" index) can then be represented as:

$$MU_i = \sum_j \alpha_j * \delta_{j,i} + \sum_k \beta_k * Ln\left(\frac{X_{i,k}}{X_{0i,k}}\right)$$

The δ is the Kronecker Delta function that equals 1.0 if the data is for country (area) "i" and its associated fixed-effect constant α_j , where "j" covers the same range as "i". Otherwise if $j \neq i$, the δ is 0.0. In this case, the estimated α picks up local issues not captured in the independent variables, such as border constraints, ethnic animosity, ties to the homeland, etc. The β are estimated parameters over "k" characteristics, as noted above. The "X" contains the value of the characteristics. The "0" term in the divisor reflects the initial condition and make the α correspond to the implied, initial fractional split (because ln(1)=0).

The normalization ("0") term could be the average or the initial value over the period, but is meant to capture the idea of change from a norm. The statistical analysis and simulation model are an attempt to understand what drives the dynamics and how changes in conditions will affect those dynamics.

The overall utility of a country is its "pull" that makes individuals want to stay where they are or for others to be attracted to the country. This utility is compared with the pull of going elsewhere. Generally speaking, the push-pull formulation combines both push (terms with negative β) and pull (terms with a positive β). The signage on β, as described here assumes that an increase in any variable is deemed a positive aspect of the areas. This relationship to signage will not always be the case.

Note that the β are only defined by characteristic and not country, because the migrants are the same individuals making the choice no matter where they are at any given moment. The analysis is an attempt to improve the understanding of what drives population-wide migration such that there is confidence in any proposed mitigation or intervention policies.

For the analysis described in this chapter, the QCT equation has the form of the multinomial logit as shown below.

$$MP_i = Exp(MU_i) / \sum_k Exp(MU_k)$$

A logit is based on a Weibull or Gumbel distributions. These are asymmetrical distributions that capture the risk aversion typically found in the choices of interest here. There is the probit version of QCT that assumes an underlying (symmetrical) normal distribution common to the product or service choices commonly analyzed in economic literature (Ben-Akiva 1985).

Typically, the QCT equation estimation assumes a mutually exclusive (binary) choice – the individual made the choice "i" or didn't; a migrant can't be in two countries at the same time. Such a regression requires maximum likelihood techniques for legitimacy. The regression generates the probability of a choice. But in a population, the probability become the fraction making that choice. Therefore the regression is continuous rather than binary and is legitimately amenable to Ordinary Least Square (OLS) regression schemes. An approximation to the binary regression that is asymptotically valid for a large number of observations is the Berkson method (Ben-Akiva 1985). This approach is inconsistent and biased for a limited number of observations under the binary assumption, but only potentially and testably inconsistent in the continuous case applied here. The discussion below derives a variant of the continuous case that uses all information available to maximize the number of independent observations.

In QCT, one of the utility functions needs to be the numeraire. In a choice set, everything is necessarily relative to the same characteristics in the other choices. The choice is arbitrary, but using the choice with the largest values of the dependent variable (in this instance, rural Mali) the impact of measurement error and small numbers is reduced.

$$\frac{MP_i}{MP_j} = Exp(MU_i)/Exp(MU_j)$$

$$\ln\left(\frac{MP_i}{MP_j}\right) = MU_i - MU_j$$

Or if "n" is the numeraire index:

$$\ln\left(\frac{MP_i}{MP_n}\right) = MU_i - MU_n$$

With the transformation of:

$$X'_{i,k} = \frac{X_{i,k}}{X_{0i,k}}$$

Then:

$$MU_i = \sum_j \alpha_j * \delta_{j,i} + \sum_k \beta_k * Ln\left(X'_{i,k}\right)$$

And

$$MU_i - MU_j = \alpha_i * \delta_{i,i} - \alpha_j * \delta_{j,j} + \sum_k \beta_k * (Ln(X'_{i,k}) - Ln(X'_{j,k}))$$

Note that if i=n,

$$\ln\left(\frac{MP_i}{MP_n}\right) = \alpha_i - \alpha_n = \ln\left(\frac{\alpha'_i}{\alpha'_n}\right) = 0$$

Where

$$\ln(\alpha'_i) = \alpha_i$$

Or

$$\ln\left(\frac{MP_n}{MP_n}\right) = \ln\left(\frac{\alpha'_n}{\alpha'_n}\right) = 0$$

By convention, α'_n is set to unity (the normalized referent value), and therefore α_n equals zero.

For illustrative purposes, let the Mali rural (R) index be I, the Mali urban (U) be 2, the neighboring countries (N) be 3, the USA (A) be 4 and the Rest-of the World (W) be 5.

While many studies just compare the numeraire to the remaining choices, separate variability among other choice pairs contain additional information on covariance.

Table 3.1 shows all the unique possibilities. In the table below, the "X" designates the inability to derive additional information for that pairing, and the "-" has redundant information with the other elements in the table. The "fraction" elements in the table represent the comparison sets where the top index (i) is compared to the other index (j) in the equations.

Table 3.1: Unique Choice Sets

	$R.1$	$U.2$	$N.3$	$A.4$	$W.5$
$R.1$	X	$2/1$	$3/1$	$4/1$	$5/1$
$U.2$	X	X	$3/2$	$4/2$	$5/2$
$N.3$	X	$-$	X	$4/3$	$5/3$
$A.4$	X	$-$	$-$	X	$5/4$
$W.5$	x	$-$	$-$	$-$	X

To apply this logic to the regression, the δ function needs to be modified. Whenever the non-numeraire ratios are estimated, the "secondary numeraire (dominator index in table 3.1) must use a -1 value for δ on that α.

This can be realized by an example. Chapter 3 contained a discussion that α_4 is really $\alpha_{4,1}$ with α_1, per the discussion of above being defined as 0.0 and α_1' being 1.0. That is

$$\alpha_{4,1} = \ln\left(\frac{\alpha_4'}{\alpha_1'}\right) = \ln(\alpha_4') - \ln(\alpha_1') = (1) * \ln(\alpha_4') - (1) * 0 = \alpha_4$$

If the regression is using the data from choice 4 compared to choice 3:

$$\alpha_{4,3} = (1) * \ln\left(\frac{\alpha_4'}{\alpha_1'}\right) + (-1) * \ln\left(\frac{\alpha_3'}{\alpha_1'}\right) = \ln(\alpha_4') - \ln(\alpha_1') - (\ln(\alpha_3') + \ln(\alpha_1'))$$
$$= \ln(\alpha_4') - \ln(\alpha_3') = \ln\left(\frac{\alpha_4'}{\alpha_3'}\right)$$

Note the negative one in the second term of the initiating equation above.

3.5. Results

Using this approach, there are 520 observations and 15 variables. The F-significance is calculated as 0.0E+00. Many of the independent variables show very high P-values, consistent with the causal BIA theory. Although p-values are problematic due to population sampling, in this instance, the analysis uses all available information (Leek & Peng 2015). The information represents the entire population and is therefore the best (even if radically inadequate) information obtainable. It is the total of information available to make any assessment or intervention decisions. Nonetheless, future work should use a Bayesian maximum-likelihood approach rather than an OLS Berkson approach to avoid potential autocorrelation problems and to better quantify uncertainty. The estimation results are shown in Table 3.2.

Table 3.2: Push-Pull Parameter Estimation

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	0	#N/A	#N/A	#N/A
WI=Income	-0.190656351	0.050678752	-3.762056925	0.0002
FA=Food availability	-0.096716122	0.028646174	-3.376231805	0.0008
GG=Govt Effectiveness	-0.167871092	0.179394349	-0.935765772	0.3498
GI=Govt Infrastructure	1.784692875	0.246148984	7.250458013	0.0000
DM=Disease Mult.	-0.229143653	0.035325973	-6.486549017	0.0000
VI=Violence Expectation	-0.031173448	0.15549478	-0.200479063	0.8412
II=Income Expectation	-0.063018841	0.062087489	-1.015000647	0.3106
FI=Food Expectation	0.46716573	0.081676371	5.719717025	0.0000
UER=LPR	-0.443127898	0.129049969	-3.433769901	0.0006
POP=POP	0.972297108	0.010790783	90.1044046	0.0000
NDI=NDI	0.002950317	0.002036594	1.448652419	0.1481
Mali Urban	-2.120287436	0.024824481	-85.41114888	0.0000
Neighbors	-2.943008918	0.02336193	-125.9745643	0.0000

USA	-12.10280441	0.025289975	-478.561346	0.0000
ROW	-5.826077215	0.022872831	-254.7160557	0.0000

The raw coefficient of determination (R^2) is 0.9994; the adjusted R^2 is 0.9974. Because the numeraire α_1 is identically 0.0, the regression is actually a Regression-through-the-origin (RTO). RTO using OLS requires special consideration of the r-square calculation. This analysis uses the modified calculation[5] noted in Eisenhauer 2003. Other studies indicate a wide range of achieved R-square, mostly in the 0.20 range, but some in the 0.97 range (Lull 2011, Mayda 2010).

Figure 3.1 indicates the normality assumption for regression is adequately valid.

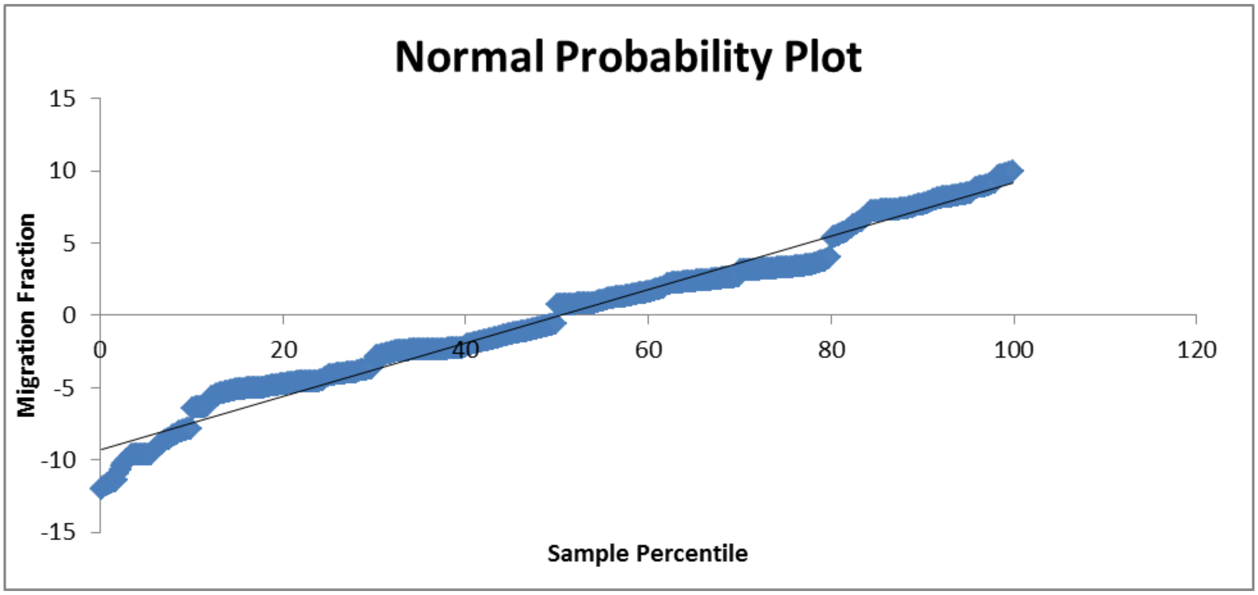


Figure 3.1: Push-Pull Normal Probability Plot

A review of the residuals (figures 3.2 and 3.3) indicates 1) no apparent heteroscedasticity; 2) appear random; and 3) have no noticeable non-linearity. Two representative plots are shown below for the unemployment rate and government effectiveness, respectively. Note that missing data was filled-in with a normalized, average, or zero value that makes the axis crossing have a high density.

[5] https://docs.oracle.com/cd/E17236_01/epm.1112/cb_statistical/frameset.htm?ch07s03s06s02.html

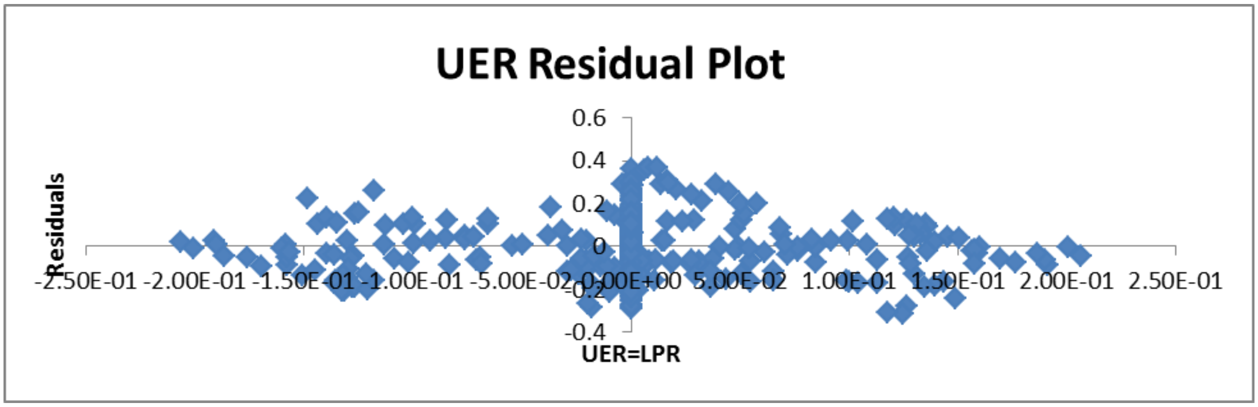


Figure 3.2: Unemployment Residual Plot

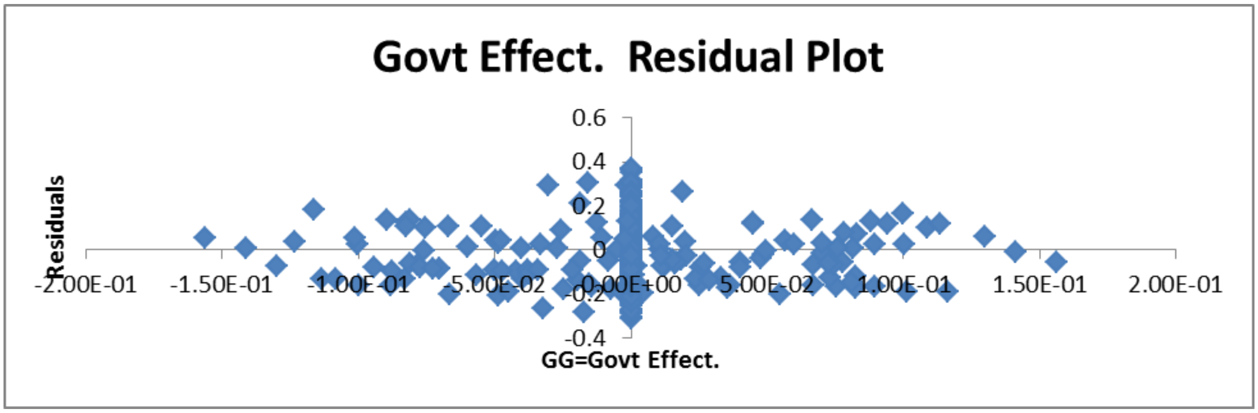


Figure 3.3: Government Effectiveness Residual Plot

Each of the estimated parameters require some interpretation. The parameters at the bottom of table 3.2 are the α terms discussed previously. The Population value is strongly positive. This feature of a support structure strongly affecting migration decisions is corroborated in other studies (Moore and Shellman 2005). Further, the parameter is near unity, raising the issue of a unit root that could invalidate the analysis in its current form. In the original MP equation, the term however is in the functional form of Exp(β*X), while the dependent variable is just a simple value. Additionally, the dependent variable is the fraction of the population in the host country, whereas, the Population term on the right-hand side of the equation is fractional population of world-wide Malians in the country associated with MP value on the left-hand side of the equation. Therefore, there is no direct cointegration (mathematical) relationship between the dependent variable and the independent population variable.

Note the GDP per Capita term is negative. The greater the GDP per capita is, the less attractive the particular country in a relative sense. That is, the pull or push of economic conditions on migration is not the same positive inducement as in rich, developed countries (Treyz 1993), but is more interwoven with clan and familial relationships (Schmeidl 1997). By any standard, Mali is an impoverished country relative to most other parts of the world. From the perspective of anyone living in the Mali homeland, an increase in local income provides added opportunity to migrate. Similarly, the food parameter is negative, suggesting a view that increased food availability means there is less harm to the remaining family if some members emigrate.

Government Effectiveness (corruption) has an apparently small negative effect. This term also applies to "captured" governments, such as North Korea. The higher the value of Government Effectiveness, the less chance there is of an individual being allowed to migrate and the less non-government "crime" there is. But, in general, the more corrupt and dictatorial the government, the less the utility of migrating to that country. Thus, there are countervailing aspects of government corruption with a net impact still being negative.

Government infrastructure has a large positive impact on the attractiveness of a country, strongly implying opportunity that residents may see as greater than those in another country.
Disease has the expected negative impacts, but is not excessively large. This result is deemed as implying that populations recognize the health issues, but are to some extent acclimated to whatever conditions exist in their locale.

Income expectation has a negative effect just like income, but to a lesser degree, indicating the short-term process of acquiring enough money "to leave." Food expectation is positive, indicating that long-term food security does have a positive effect on well-being and the desirability of being in a particular area. Given that subsistence farming is a dominant occupation in Western Africa, food supply is an important part of daily life.
Increased unemployment has a negative effect on the utility of a country as a host nation, as one would expect.

Neither natural disasters nor violence appears to have a significant effect on migration. The model was tested using both actual and expected values. There was no noticeable difference or impact. Other researchers have also noticed the lack of relationship of long-term migration with natural disasters (Bohra-Mishra et al. 2014). It would appear that violence and natural disasters are associated with phenomena separate from long-term migration, possibly affecting only a short-term forced migration. Short-term migration is the subject of the Chapter 4.

3.6. Summary

The use of the BIA approach appears to offer an improved means to estimate the parameters used in simulation models of migration. It captures both the push and pull characteristic affecting migration decisions. It additionally provides a statistically outstanding appraisal of the migration choices across a set of interdependent counties. The approach suggests the benefit of using "perceptions" of conditions (such as those associated with governance), rather than just directly measured impacts. Further, the modeling omits variables that do not act as the information an individual actually uses to make permanent migration choices. The design empowers a more causal consideration of the climate impacts on migration by including economic and productivity effects explicitly within the simulation model, and in having tangible outcomes (information actually available to the individuals) be part of the decision calculus.

Many observers consider the Arab spring (Johnstone & Mazo 2011), the Sudanese unrest (Boslough, Backus, et al., Hendrix and Glaser 2007, Nordas & Gleditsch 2007), and the war in

Syria (Gleick 2014, Brown & Crawford 2009) to be the result of changing climate conditions (mostly drought in these instances). When the results of this analysis are incorporated into the simulation model discussed later in this report, the analysis suggests that the current concerns with the massive flux of migrants, for example, into Western Europe, are but a precursor to much larger and more geographically diverse movement of populations. In most cases the pull will necessarily be to the more developed nations.

3.7. References

Backus, G. A., & Glass, R. J., 2006. An agent-based model component to a framework for the analysis of terrorist-group dynamics. *Sandia Report, SAND2006-0860*, Sandia National Laboratories, Albuquerque, NM

Backus, G., M. Bernard, S. Verzi, A. Bier, and M. Glickman, 2010. .Foundations to the unified psycho-cognitive engine. Technical Report SAND2010-6974, Sandia National Laboratories.

Ben-Akiva, Moshe E., and Steven R. Lerman, 1985. *Discrete choice analysis: theory and application to travel demand*. Vol. 9. MIT press.

Bernard, M. L., Backus, G. A., Bier, A. B., & Branch, S. F. , (2014). Behavioral Influence Assessment (BIA): A multi-scale system to assess dynamic behaviors within groups and societies across time. *Advances in Cross-Cultural Decision Making5*, 161.

Bohra-Mishra, Pratikshya, Michael Oppenheimer, and Solomon M. Hsiang, 2014. "Nonlinear permanent migration response to climatic variations but minimal response to disasters." *Proceedings of the National Academy of Sciences*: 201317166.

Boslough M, Sprigg J, Backus G, Taylor M, McNamara L, Fujii J, Murphy K, Malczynski L, 2004. Climate change effects on international stability: a white paper. Sandia National Laboratories, Albuquerque.

Brown, Oli, and Alec Crawford, 2009. "Rising temperatures, rising tensions: climate change and the risk of violent conflict in the Middle East."

Coniglio, Nicola D., and Giovanni Pesce, 2014. "Climate Variability and International Migration: an empirical analysis." *Environment and Development Economics*: 1-35.

Eisenhauer, J. G., 2003. Regression through the origin. *Teaching Statistics*,*25*(3), 76-80. https://online.stat.psu.edu/~ajw13/stat501/SpecialTopics/Reg_thru_origin.pdf

Francis, D. J., 2013. The regional impact of the armed conflict and French intervention in Mali. *Norwegian Peacebuilding Resource Centre*. http://peacebuilding.no/var/ezflow_site/storage/original/application/f18726c3338e39049bd4d554d4a22c36.pdf

Gleick, Peter H., 2014. "Water, drought, climate change, and conflict in Syria." Weather, Climate, and Society 6, no. 3) : 331-340.

Hendrix, Cullen S., and Sarah M. Glaser, 2007. "Trends and triggers: Climate, climate change and civil conflict in Sub-Saharan Africa." Political geography 26, no. 6: 695-715.

Johnstone, Sarah, and Jeffrey Mazo, 2011. "Global warming and the Arab Spring." Survival 53, no. 2: 11-17.

Leek, J. T., & Peng, R. D., 2015. Statistics: P values are just the tip of the iceberg. *Nature*, *520*(7549), 612-612.

Llull, J., 2011. Understanding International Migration: Evidence from a New Dataset of Bilateral Stocks (1960-2000). *Manuscript Universitat Autonoma Barcelona*.

Lo, A., Chernoff, H., Zheng, T., & Lo, S. H. , 2015. Why significant variables aren't automatically good predictors. *Proceedings of the National Academy of Sciences*, 201518285.

Mayda, Anna Maria, 2010. "International migration: A panel data analysis of the determinants of bilateral flows." *Journal of Population Economics* 23, no. 4: 1249-1274.

McFadden, D., 1986. "Econometric Model of Probabilistic Choice," in Structural Analysis of Discrete data with Econometric Applications, ed. C.F. Manski and D. McFadden, Cambridge, MA, MIT Press.

McFadden, D., 1974. "Conditional Logit Analysis of Qualitative Choice Behavior," in Frontiers in Econometrics, Ed. P. Zarembka, New York, Academic Press.

McFadden, D., 1982. "Qualitative Response Models," in Advances in Econometrics, Ed. Werner Hildenbrand, Cambridge University Press, New York.

Moore, Will H., and Stephen M. Shellman, 2007. "Whither will they go? A global study of refugees' destinations, 1965–1995." *International Studies Quarterly* 51, no. 4: 811-834.

Nordås, Ragnhild, and Nils Petter Gleditsch, 2007. "Climate change and conflict." Political Geography 26, no. 6 : 627-638.

Treyz, George I., Dan S. Rickman, Gary L. Hunt, and Michael J. Greenwood, 1993. "The dynamics of US internal migration." *The review of Economics and Statistics*: 209-214.

4. SHORT-TERM MIGRATION DYNAMICS

In the previous chapter, it was clear that violence and natural disasters appeared to have little or no statistical effect on long-term (permanent) migration. Yet, the immense Dadaab refugee camp in Kenya is strongly associated with violence in Somalia.[6] The 2015 migration to Europe is strongly associated with the war in Syria.[7] The recent floods in Myanmar[8] make the news for the number of displaced persons involved, as do the earthquakes in Nepal.[9] Short-term, migration, also called temporary or forced migration, might be best considered a separate phenomenon from long-term or permanent migration. Additionally, natural and manmade disasters are largely implicated as the cause of short-term migration dynamics.

4.1. Basic Concepts

The approach is to again test a BIA construct. In this case, the issue is what fraction of the population chooses to leave the area because of extreme events, either from a natural disaster or a military conflict. In the context of the previous chapters, the migration is a push phenomenon. The analysis here considers the affected population from the extreme event and the perceived ability of the government to provide necessary services or to manage the situation. It also considers the individual resources to accommodate the event as measured by GDP per capital. The GPD per capita, captures two aspects: the general wealth of the area plus the infrastructure it implies for independently managing the situation. These considerations determine the relative utility of going to a camp versus staying in-place. Although there are diminishing returns for the ability of government or individual resources to handle a situation as it becomes more catastrophic, the analysis here assumes a linear, albeit composite, relationship among all the variables in the utility function. The composite feature is meant to denote that a single term of the linear utility function may be composed of multiplicatively combined elements, such as those noted above. A theoretical and more complicated formulation of governmental response and GDPPC impacts would not have allowed a practical means for agencies to estimate parameters and confidence.

There is little research on the topic of predicting the number of displaced persons. The purpose here is to determine confidence intervals which relief and aid agencies can use to estimate required resources to handle the situation. These resources can take on two aspects. One aspect could help agencies determine proactively what is needed in reserve to manage future event. The second aspect is help agencies respond reactively to events as they happen. In the former instance, as supported by the data, local governments can be trained and equipped to mitigate the number of individuals feeling they must leave to maintain security.

---

[6] http://www.theguardian.com/global-development/2015/apr/17/dadaab-refugee-camp-closure-risk-350000-somali-lives-amnesty
[7] http://www.cfr.org/migration/europes-migration-crisis/p32874
[8] http://www.rfa.org/english/news/myanmar/aid-08052015174241.html
[9] http://www.aljazeera.com/news/2015/07/nepal-earthquake-150725054637088.html

4.2. Data

The analysis uses data from the Internal Displacement Monitoring Centre (IDMC[10]) and United Nations High Commissioner for Refugees (UNHCR[11]) to account for the number of internally displaced persons (IDP), refugees, and those in a similar situation. It uses the World Bank WDI data noted in the previous chapter for the total population of the country and for GDP per capital (real $US). Conflict data comes from the Uppsala Conflict Data Program (UCDP[12]) database. The analysis uses the information on total number of military and civilian causalities. The total number of affected persons from natural disasters comes from the Centre for Research on the Epidemiology of Disasters (CRED[13]). Information on government effectiveness comes from the World Banks WGI data discussed in the previous chapter.

The table 4.1 below provides more data on these resources. For the purposes here, there is only adequate data for the years 2000 through 2014. The data in this case, covers the 56 (officially recognized and unrecognized) countries within Africa with 692 observations.

Table 4.1: Short-term Migration Data Sources

Data	Dates available	Source
Economic and Social indicators	1960 - 2013	World Development Indicators (World Bank) http://data.worldbank.org/products/wdi
Worldwide Governance Indicators	1996 - 2013	http://databank.worldbank.org/data/views/reports/tableview.aspx
Refugee and Internally Displaced Persons	2000 - 2013	UNHCR (The UN Refugee Agency) www.unhcr.org/
Disasters Natural, Climatological, Geophysical, Meteorological, etc.	1900 - present	EM-DAT, The International Disaster Database, Centre for Research on the Epidemiology of Disasters – CRED http://www.emdat.be/database
Conflict	1989 - 2013	Uppsala Conflict Data Program (Date of retrieval: yy/mm/dd) UCDP Conflict Encyclopedia: www.ucdp.uu.se/database, Uppsala University Department of Peace and Conflict Research conflictdatabase@pcr.uu.se

4.3. Mathematical Construct

In the case of short-term migration, the choice is to stay or go, or equivalently to remain in the camp or return. The analysis here models the fraction of the population that would be in a camp after an extreme event.

---

[10] http://www.internal-displacement.org/global-figures
[11] http://www.unhcr.org/pages/49c3646c4d6.html
[12] http://www.pcr.uu.se/research/UCDP/
[13] http://www.emdat.be/database

The choice logic follows that of the previous chapter. The migration population (MP) fraction is.

$$MP_i = Exp(MU_i)/\sum_k Exp(MU_k)$$

Even though this analysis treats the choice as a continuous probability and translates the result to a fraction of the population moving, there are still only two choices represented: stay (s) or go (g). The choice equation is then:

$$MP_g = Exp(MU_g)/(Exp(MU_s) + Exp(MU_g))$$

where MU is the utility. Dividing the top and bottom of the MP equation by $Exp(MU_g)$ results in:

$$MP_g = 1/(1 + (Exp(MU_g - MU_s))$$

It is not possible to separate the utility of going versus staying because, as noted in the previous chapters, one of the choices has to act as the numeraire for the other. The utility function, at a theoretical level, typically takes on a linear or log-linear formulation:

$$MU_i = \alpha_i + \sum_k \beta_k * X_{i,k}$$

or

$$MU_i = \alpha_i + \sum_k \beta_k * Ln\ (X_{i,k})$$

The change in choice is due to a change in condition. The “normal” or normed condition is defined as the absence of a natural disaster or a military conflict. As a convention, the α of the numeraire is set to 0.0, hence, the estimate choice equation is simply:

$$MP_g = 1/(1 + (Exp(-MU_s))$$

Or by remembering the α and β would actually have a reversed sign:

$$MP_g = 1/(1 + (Exp(MU_s))$$

This approach is consistent with other studies (Joarder and Miller 2012), but using an aggregate population versus individual evaluation (Williams et al. 2010).

The equation is linearized by making a transformed variable:

$$\varphi_i = \ln(\frac{1}{MP_g} - 1)$$

such that

$$\varphi_i = MU_i$$

where MU_i is the country dependent utility to leave the function form used in the previous chapter:

$$MU_i = \sum_j \alpha_j * \delta_{j,i} + \sum_k \beta_k * Ln(X'_{i,k})$$

The α_j capture the local cultural or institutional traditions in each country that may affect the propensity to leave the area, or they can reflect the self-reliance aspects of the population. As in the beginning sections of the previous chapter, the δ is 1.0 only if i=j; otherwise it is 0.0. There will be 56 α, one for each country. The β are not by country because of our interest in determining if there are universal decision characteristics that allow their generic use for anticipating and responding to disasters, no matter where they occur or whether there is even previous data on disaster impacts. The analysis tested many linear and log-linear combinations of military events, natural disaster events, government effectiveness, and GDP per capita. These mostly produced parameterizations with very low predictive power, having R-squares barely above zero and usually below 0.25. Such results are common in the literature (Saldaña-Zorrilla 2009.

Some studies that use a very large number of independent variables have produced R-square near 0.8, but the authors note their lack of predictive power (Henry et al. 2003, Claydon 2013).

For robustness and simplicity, and statistical confidence, this analysis ultimately focused on only two (composite) dependent variables: One associated with natural disasters and one with military conflicts. The propensity to flee should be affected directly by the fraction of the population affected by the event, but inversely affected by the confidence in the local government to manage the event and inversely affected by the local economic (resilience and infrastructure) conditions. For economic resilience, the analysis uses GDP per capita (GDPPC). The perception of government confidence is more important than the reality. Therefore the analysis uses the World Bank governance indicators for the country's Government Efficiency $\theta_{d,i}$ for natural disasters, and Political Stability $\theta_{c,j}$ for military conflict.

The fractional population affected by a disaster only covers the range from 0.0 to 1.0. For the exponential formulation of the QCT to work effectively without uninterpretable α and β, the affected population fraction needs to cover the range from positive infinity to negative infinity. If the fraction for a country is "$f_{j,i}$", a transformation that meets these criteria is:

$$\varphi_{j,i} = \ln(\frac{1}{f_{j,i}} - 1)$$

where the "j" is a d" for disaster and a "c' for conflict.

As noted in the previous chapter, the use of the logarithm "ln" implies the perception as being a relative change in conditions. The term within the ln can be rewritten as

$$f/(1-f)$$

If one recognized that a negative sign in front of the logarithm inverts the argument, this term is now an odds ratio of those affected to those not affected. This is typically a measure of risk, for example in disease management (Schechtman 2002). Thus, the transformation has a causal basis. Another benefit it that it normalizes the data. The recorded fraction of the population affected is typically very small, seldom rising beyond 10%. The data are highly skewed near the origin, whereas after the transformation, the data closely approximates a normal distribution.

There are additional nuances to using the data. An initiating event happens over a span of minutes to days. The refugee or displacement experience can be an extended experience. The UNHCR data indicates a roughly 2 year period of displacement. The analyst only has annual data and resolution, and therefore can only meaningfully use integer time periods. With 3 years clearly being too long, and 1 year having no effect, a 2 year exponential filter generates an adequate approximation to the remembered event for the purposes here. This lagging means that involved individuals may only experience the physical event momentarily, but the perceived impact of it lingers, either because they are afraid to go back, or it takes time for the support structure "back home" to return to normal. The use of the lag structure on the reported number of individuals affected by an initiating event dramatically improved the alignment with the record number in the camps.

The BIA approach suggests that individuals compare normal conditions to current events for making decisions. Thus, individuals living in an area prone to natural disasters, for example, flooding, should be acclimated to significant levels of flooding. The analysis did not show this comparative process to be significant or at least active. This suggests that people feel insecure whenever there is a threat and even "routine" threats contain random elements that lead to causalities or intense damage. Hence, rather than learning to adapt to the environments, the adaptive response is to leave as the danger becomes evident.

As a last point on data reconciliation, the relief agencies seldom have the resources to fully understand the number of people affected in, say, a remote region. Further small local events such as fires or landslides displace a small number of people who may never show up on any roster. There is a high degree of measurement error in the data, and natural background noise (criminal conflicts or natural disasters) means that rather than assuming a 0.0 in the refugee, disaster, or conflict data actually means no displaced population, it might be better to assume a very small number such as one millionth of the population. This adjustment adds robustness and is also a necessity of prevent division by zero or taking the logarithm of 0.0 in the regression equation.

By noting the lagged transformed variables as$L\varphi_{j,i}$, the utility function of going (g) for the two independent variables used in analysis can be written as:

$$\varphi_{g,i} = \sum_j \alpha_j * \delta_{j,i} + \beta * \frac{\varphi_{d,i}}{L\theta_{d,i} * GDPPC_i} + \gamma * \frac{\varphi_{c,i}}{L\theta_{c,i} * GDPPC_i}$$

4.4. Results

The analysis produces an adjusted R-square of approximately 0.88. It indicates that both development and government-training aid can have a significant impact. Similar R-square values are obtained by only using the $\varphi_{j,i}$ terms without the denominator adjustments. Because of the short time period over which there is sufficient data, the α readily picks up the local average GDPPC variation across countries. Because of the correlation between governance indicators within a country, the α can also pick up a significant part of that variation. In that the above equation better parses the causal factors, leaving the α to only capture un-modelled effects, the value using the equation above becomes more apparent. Aid agencies and governments can never have adequate information on where the next disaster will occur or of its magnitude. There is always a large amount of uncertainty, but that uncertainty only affects 13% of the outcome variation, with 87% being captured by using the limited information that is reasonably available. The agencies can be an aware of a country's preparedness and help promote that preparedness. The β and γ are intended to be an unsophisticated but generic impact dynamic that enables an immediate and useful estimate of potential displacement from early accounts of the disasters. By using the 95% confidence value of the parameters, there can be an upper level of confidence on the preparation required to accommodate the event.

One of the analysis tests was to treat existence of a natural disaster or a conflict as a binary variable; a natural disaster or military conflict simple existed or did not. Nearly 20% of the consequences or impacts are recognizable by that knowledge alone. This feature might be useful in the absence of any quantified information.

The estimation of the previous equation produces a substantial F-static indicating significance well below the 0.001 level. The equation fits the data as a linear model, per figure 4.1.

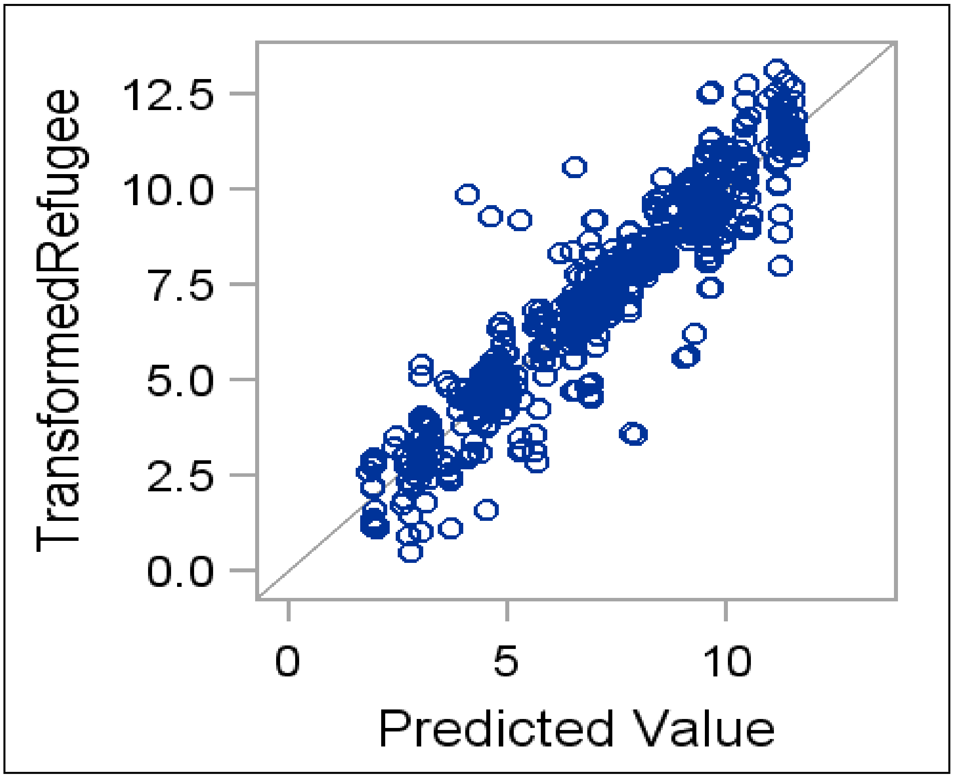


Figure 4.1: Linear Probability Plot

The parameter estimates are shown below (table 4.2). Five countries did not have sufficient data to make a meaningful estimate of local (fixed) effects – which had the secondary effect of reducing the degrees of freedom for the regression.

Table 4.2: Short-term Migration Parameter Estimation

Parameter Estimates						
Variable	**Label**	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	Intercept	1	6.48134	0.25790	25.13	<.0001
TransDeath/(PolStab*GDP/capita)	TransDeath/(PolStab*GDP/capita)	1	-44.95207	8.29545	-5.42	<.0001
TransDis/GovEff*GDP/capita	TransDis/GovEff*GDP/capita	1	21.14567	18.69098	1.13	0.2584
algeria	Algeria	1	1.67449	0.36098	4.64	<.0001
angola	Angola	1	-2.14739	0.36048	-5.96	<.0001
benin	Benin	1	3.20284	0.35939	8.91	<.0001
botswana	Botswana	1	3.99884	0.36187	11.05	<.0001
burkina_faso	Burkina Faso	1	3.02924	0.35941	8.43	<.0001
burundi	Burundi	1	-2.64319	0.41063	-6.44	<.0001
cabo_verde	Cabo Verde	0	0	.	.	.
cameroon	Cameroon	1	0.90557	0.36205	2.50	0.0126
central_african_republic	Central African Republic	1	-2.42861	0.36614	-6.63	<.0001
chad	Chad	1	-1.66122	0.36163	-4.59	<.0001
comoros	Comoros	1	1.33099	0.36066	3.69	0.0002

Parameter Estimates						
Variable	**Label**	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
congo__dem__rep_	Congo, Dem. Rep.	0	0	.	.	.
congo__rep_	Congo, Rep.	1	-1.54870	0.36021	-4.30	<.0001
cote_d_ivoire	Cote d'Ivoire	1	-1.36433	0.36024	-3.79	0.0002
djibouti	Djibouti	1	0.44328	0.36144	1.23	0.2205
egypt__arab_rep_	Egypt, Arab Rep.	0	0	.	.	.
equatorial_guinea	Equatorial Guinea	1	0.76131	0.36201	2.10	0.0359
eritrea	Eritrea	1	-3.48964	0.36145	-9.65	<.0001
ethiopia	Ethiopia	1	1.18191	0.40419	2.92	0.0036
gabon	Gabon	1	2.82000	0.36181	7.79	<.0001
gambia__the	Gambia, The	1	0.00655	0.35941	0.02	0.9855
ghana	Ghana	1	0.72576	0.35983	2.02	0.0441
guinea	Guinea	1	0.60325	0.36926	1.63	0.1028
guinea_bissau	Guinea-Bissau	1	0.39199	0.36004	1.09	0.2767
kenya	Kenya	1	0.74677	0.36199	2.06	0.0395
lesotho	Lesotho	1	4.77860	0.36110	13.23	<.0001
liberia	Liberia	1	-3.27485	0.36727	-8.92	<.0001
libya	Libya	1	-2.04293	0.60632	-3.37	0.0008
madagascar	Madagascar	1	5.01181	0.35980	13.93	<.0001
malawi	Malawi	1	3.06784	0.36072	8.50	<.0001
mali	Mali	1	1.80449	0.35961	5.02	<.0001
mauritania	Mauritania	1	-1.96376	0.36055	-5.45	<.0001
mauritius	Mauritius	1	3.26955	0.36187	9.04	<.0001
morocco	Morocco	1	2.95158	0.36071	8.18	<.0001
mozambique	Mozambique	1	4.64199	0.35936	12.92	<.0001
namibia	Namibia	1	0.76765	0.36164	2.12	0.0342
niger	Niger	1	2.70227	0.36006	7.50	<.0001
nigeria	Nigeria	1	2.40814	0.36307	6.63	<.0001
rwanda	Rwanda	1	-1.73581	0.36306	-4.78	<.0001
sao_tome_and_principe	Sao Tome and Principe	1	1.99372	0.36204	5.51	<.0001
senegal	Senegal	1	0.25639	0.36033	0.71	0.4770
seychelles	Seychelles	1	0.98353	0.36201	2.72	0.0068
sierra_leone	Sierra Leone	1	-1.71391	0.36580	-4.69	<.0001
somalia	Somalia	0	0	.	.	.
south_africa	South Africa	1	4.91993	0.36154	13.61	<.0001
south_sudan	South Sudan	0	0	.	.	.
sudan	Sudan	1	-3.01781	0.36480	-8.27	<.0001
swaziland	Swaziland	1	3.49078	0.36177	9.65	<.0001
tanzania	Tanzania	1	2.53686	0.36017	7.04	<.0001
togo	Togo	1	-0.69745	0.35943	-1.94	0.0528
tunisia	Tunisia	1	1.70601	0.36184	4.71	<.0001
uganda	Uganda	1	-0.41682	0.36659	-1.14	0.2560
zambia	Zambia	1	3.86002	0.35945	10.74	<.0001

An explicit Zimbabwe coefficient was left out of the regression to avoid regression-through-the-origin, as discussed in the previous chapters. The intercept term is then actually the fixed effect constant for Zimbabwe.

4.5. Summary

The purpose of this exercise was to determine a useful, although not necessarily precise, means of estimating the extent of forced migration due to natural disasters (such as from climate change) or military conflict (climate change being an underlying factor.)   Large p-values are best interpreted as underlying uncertainty, rather than as a loss of predictive power. They indicate the inherent and realistic inability to precisely predict, but do not prevent the adequate quantification of the risks that affect aid decisions.

4.6. References

Claydon, KJ,2012. A global model of human migration. Centre for Advanced Spatial Analysis, University College. London. July, 2012. *https://www.bartlett.ucl.ac.uk/casa/pdf/paper186*

Henry, Sabine, Paul Boyle, and Eric F. Lambin, 2003. "Modelling inter-provincial migration in Burkina Faso, West Africa: the role of socio-demographic and environmental factors." *Applied Geography* 23, no. 2: 115-136.

Joarder, Mohammad Abdul Munim, and Paul W. Miller, 2013. "Factors affecting whether environmental migration is temporary or permanent: Evidence from Bangladesh." *Global Environmental Change* 23, no. 6: 1511-1524.

Saldaña-Zorrilla, Sergio O., and Krister Sandberg. "Impact of climate-related disasters on human migration in Mexico: a spatial model." *Climatic change* 96, no. 1-2 (2009): 97-118.

Schechtman, Edna, 2002. "Odds ratio, relative risk, absolute risk reduction, and the number needed to treat—which of these should we use?" *Value in Health* 5, no. 5: 431-436.

Williams, Nathalie E., Dirgha J. Ghimire, William G. Axinn, E. Jennings, and M. Pradhan, 2010. "A micro-level approach to investigating armed conflict and population responses." Population Studies Center, University of Michigan.

5. SPILLOVER VIOLENCE

The last choice process included in the simulation model is the dynamics of spillover violence. Does violence in one area lead to violence in another, due to the expansion of economic hardship or government action/inaction? Because this study focuses on Mali, there are a limited number of observations for establishing a high degree of confidence in answering this question. The methods developed, however, can be applied to other areas to confirm or invalidate the phenomena described in this analysis.

5.1. Basic Concepts

This analysis addresses the potential for the spillover of violence in Mali to neighboring countries and vice versa. The analysis considers the relative level of violence in a neighboring country as it affects violence in the country of interest. An assumption is that a moment of violence is not a concern, but enduring (perceived) violence in another country may affect the acceptability of violence as a means to either address grievances in the country of interest (COI) or as a defense against parties in the neighboring country using the COI as safe harbor or as a launch point for violence in their country of origin.

The key element is the comparison of the perceived level of violence across the border relative to the "normal" levels of violence within the broader of the COI. Violence is measured using the World Bank WGI Rule of Law index, as noted in the chapter on Long-term Migration Dynamics. Previously, the violence index was scaled to a range between 0 and 100. Here the raw WGI index is scaled to span 0 and 100%. That is, the index is converted to a fraction for use as the dependent variable in a QCT formulation. This approach implicitly assumes that if the violence index is at its maximum possible value, 100% of the population perceives they are affected by the violence.

Income levels are part of the analysis, with low income levels being a proxy for both grievance and for the infrastructure or resilience the community has to combat spillover violence. Unemployment of males between 15 and 24 (using the labor participation rates as a proxy to estimated unemployment) is included as a means to capture discontent and the absence of productive options for the male population.
Food availability, as in the long- term migration analysis of Chapter 3, was used to assess how food shortages affect grievances and potential conflict. Lastly, the analysis includes the perception of government effectiveness. This term uses the World Bank WGI Control of Corruption index as a proxy.

At the time of the analysis, data was only available for the years 1998 through 2012 and constitutes 42 observations.

5.2. Mathematical Construct

Per the previous chapter, the fractional population affected by violence (VP) is specified using QCT:

$$VP_i = 1/(1 + Exp(-VU_i))$$

with utility of Violence (VU) being specified as:

$$VU_i = \alpha_i + \sum_{j=u,r,n} \omega_{ji} * Ln\left(\frac{AVP_j}{VP_{0i}}\right) + \omega_{4i} * Ln\left(\frac{WI_i}{WI_{0i}}\right) + \omega_{5i} * Ln\left(\frac{UER_i}{UER_{0i}}\right) + \omega_{6i} * Ln\left(\frac{GG_i}{GG_{0i}}\right)$$

In this instance, the specific indices designated as u, r, and n represent Mali Urban, Mali rural, and, Mali (border-sharing) neighbors, respectively. The “0” index represents initial values. Note in the above equation that the “j” represents the neighboring country, and “i” represents the country of interest. Logarithms are again used to denote relative change.

5.3. Results

The unadjusted R-square for the analysis is 0.97 (table 5.1). The adjusted R-square of only 0.81, due to the limited data used in this analysis. The p-values are very poor, but the sign on many of the parameters is fairly robust.

Table 5.1: Spillover Violence Parameter Estimation

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	0	#N/A	#N/A	#N/A
AVP=Violence	0.106143707	0.10535685	1.007468488	0.325177922
WI=Income	-0.012565092	0.045579388	-0.275674866	0.785492289
UER=LPR	0.044826069	0.118700149	0.377641217	0.709487016
GG=Govt Effect.	-0.259484147	0.199445751	-1.301026195	0.207344471
food	-1.55819802	0.160713821	-9.69548236	3.32046E-09
MR	0.410647892	0.018284104	22.45928447	3.64523E-16
NE	0.606267212	0.018429125	32.89723247	1.4955E-19

From the regression, it seems reasonable that an environment of violence in an one area increases the potential for violence in a neighboring area. Income appears to have a minor and highly ambiguous effect, with even the signage of the impact unresolved. At low income levels, increased income could be due to illicit activities or used to purchase weapons, as well as for the opposite purposes to promote well-being and a desire to avoid violence.

Similarly, increased unemployment appears to potentially increase violence, but the uncertainty of the parameter signifies a significant probability of having a reversed sign. Possibly very low levels of income indicate helpless poverty and the inability to participate in extended violence.

The government effectiveness (Control of Corruption) has some interesting implications. It is likely positive and impactful. This parameterization suggests that the greater the government effectiveness (less corrupt), the greater the likelihood of violence! For the countries in question, governance is at low levels. This typically means the government is very corrupt and may have strong police state tendencies. Police states such as North Korea are very safe. The governments have a true monopoly on violence and abundant control. Iraq stayed relatively calm under the control of strong dictator. Many other examples abound in Africa itself. Consequently, the counterintuitive sign of the government effectiveness parameter makes sense. This result of the analysis is likely meaningful.

Similarly, the food parametrization is not likely real and meaningful. It indicates the more food there is, the less the violence. This suggests that the lack of food security may be a source of grievance empowering violence.

The basis for this analysis is the implicit assumption that grievances are a source of violence. Other researchers have argued that resource predation is required to support and justify the costs of executing violence. Such analysis shows r-square results on the order of 0.25 but the results are much more robust than those presented here (Collier and Hoeffler2004). These types of analyses assume government participation in the conflict, whereas the assumption of this work is that the grievances are the foundation of the violence and that governments merely respond to the violence or act as an inhibitor of it (i.e., essentially a competitor to it). Nonetheless, a literature review indicates the work here is the first to address the spatial spillover dynamics of conflict. Previous work tends to exclusively focus on the *economic* spillover of conflicts in neighboring areas. (DeGrout 2010, Murdoch and Sandlet 2002)

5.4. References

Collier, P., & Hoeffler, A., 2004. Greed and grievance in civil war. *Oxford economic papers*, *56*(4), 563-595.

De Groot, Olaf J., 2010. "The spillover effects of conflict on economic growth in neighbouring countries in Africa." *Defence and Peace Economics* 21, no. 2: 149-164.

Murdoch, James C., and Todd Sandler, 2002. "Economic growth, civil wars, and spatial spillovers." *Journal of conflict resolution* 46, no. 1: 91-110.

6. A REGIONAL MODEL OF HUMAN MIGRATION AND CLIMATE CHANGE EFFECTS

Asmeret Naugle, George A. Backus, Vincent C. Tidwell,
Elizabeth Kistin-Keller, Daniel Villa

6.1. Abstract

Changing climatic conditions are likely to place increased stress on vulnerable populations through intensifying damage to homes and critical infrastructure, reduced food production, compromised health and hygiene, and land and environmental degradation. Deteriorating conditions can increase pressure on affected populations to migrate within and beyond their countries. Spillover effects from increased migration may include conflict, crime, terrorism, economic disruptions, and increased demand for humanitarian aid. Given the complexity of factors influencing human migration, quantitative tools are needed to identify populations at risk and explore policy strategies and adaptive measures. Toward this need, a system dynamics-based model is developed that couples migration behavior with the interacting dynamics of economy, labor, population, violence, governance, water, food, and disease. A regional model focused on Mali in western Africa has been adopted for the first test case.

6.2. Introduction

Some 215 million people, or 3 percent of the world's population, are believed to live outside their countries of birth (United Nations 2009a), while millions more have migrated internally (United Nations 2009b). Migration is often undertaken to improve one's quality of life, in response to such factors as the economy/employment (Massey et al. 1993), land degradation (Ghimire and Mohai 2005), social networks (Davis et al. 2013), and community factors (Grote et al. 2006). In other cases migration is forced, driven by conflict (IDMC 2011; Salehyn and Gleditsch 2006), environmental disasters (El-Hinnawi 1985), or other influences. Migration is a decision that impacts the welfare of the household, the home community, and in the end the whole economy (Azam and Gubert 2006). Spillover effects of increased migration may include greater need for costly humanitarian aid, increased crime and terrorism, interruption of international trade, and mobilization of peace keeping forces (CAN 2009; Defense Science Board 2011; National Intelligence Council 2012).

Growing concern over climate change has drawn recent attention to human migration, suggesting that intensifying floods, droughts, and sea level rise could result in unprecedented migration (Myers 2002). Projections of environmentally induced migration vary widely, from 200 million (Brown 2008) to 700 million (Christian Aid 2007) by 2050. While many may question these numbers (Black 1998; 2001), few disagree with the fact that related environmental challenges will put increased stress on at-risk populations and may motivate internal and international migration.

Homer-Dixon (1991; 1994) gives case study evidence that links population growth, environmental deterioration and political violence to migration. Reuveny and Moore (2009) found deteriorating environmental conditions in a developing country promotes out-migration to the developed world, all else being equal. In a study of Nepal's Chitwan Valley during the late 1990s environmental deterioration (e.g., declining land cover, increasing population density, perceived declines in agricultural productivity) was found to lead to short-distance moves within the immediate vicinity (Massey et al. 2007). Similarly, analysis of survey data for both migrants and non-migrants in 12 countries suggest that while long-term environmental events, such as droughts, have no significant effect on internal migration, sudden-onset environmental events in the form of floods significantly increase the likelihood of migration. Furthermore, individual perceptions of negative environmental conditions can motivate people to move. They also found that people tend to respond to long-term environmental problems, such as environmental degradation, with adaptation rather than migration. Ultimately, migration dynamics are complicated, with environmental conditions being one of many factors that mutually influence migration (Wood 2001).

Analytical tools are needed to assist in identifying populations at greatest risk and exploring robust policy strategies and adaptive measures (UN Framework Convention on Climate Change 2013). A variety of approaches have been taken toward analyzing environmental impacts on migration trends. Perch-Nielsen (2004) argued that migration induced by climatic hazards has not been integrated into migration models; therefore, existing climate and migration models cannot be simply linked. She proposed four conceptual models linking climate change and migration, addressing sea level rise, floods, tropical cyclones and drought. Other qualitative/conceptual models have been developed as the result of a variety of environmentally induced migration case studies (McLeman and Smit 2006; Gilbert and McLeman 2010; Black et al. 2011). Empirically based models have been developed through regression of large sets of data generally structured around a conceptual model of migration; for example, informal cost-benefit (Reuveny and Moore 2009) and the gravity model (Afifi and Warner 2008). Economic modeling has also been pursued; specifically, a general equilibrium (microeconomic) model of environmental migration (Siyaranamual 2009; Chichilnisky and DiMatteo 1998). Agent based modeling provides a framework to simulate human behavior, allowing agents (individuals or groups of individuals) to interact with the environment in complex ways including environmentally motivated migration. Examples include Mena et al. (2011) who developed an agent-based model to simulate deforestation change associated with land use patterns of frontier migrant framers in Northern Ecuadorian Amazon. Kniveton et al. (2012) used an agent-based model developed around the theory of planned behavior to explore how climate and demographic change combine to influence migration within and from Burkina Faso.

Here a model of human migration is developed with the unique feature that migration behavior is integrated and tightly coupled with the dynamics of the regional economy, labor, population, violence, governance, water, food, and disease. The model is formulated within system dynamics architecture. The focus of this paper is a prototype model developed to explore migration dynamics for developing nations in western Africa. Specifically, Mali serves as the focus of the analysis, with migration considered within country, between rural and urban, as well as internationally, including migration to and from neighboring countries and more distant developed nations. The model provides a quantitative means of exploring the effects of climate change on social, economic, infrastructure and resource dynamics/constraints and their interdependent influence on human migration. The model is also designed to explore alternative future adaptation pathways to understand the efficacy and robustness of alternative policy strategies, determining what pre-emptive adaptive measures are most necessary when and where.

6.3. Methods

To assess likely dynamics of climate-induced migration, we created and simulated a system dynamics model of movement of Malian people in response to climate change and related variables. The model specifies causal relationships between economic, climate, population, and decision-making factors. The model has a 70-year time horizon, beginning in 1990 and projecting out to 2060. It was created using Vensim (Ventana Systems 2013) and Microsoft Excel software, and is specified using difference equations and simulated with Euler integration. A broad-scale overview of which model sectors affect each other is shown in figure 6.1.

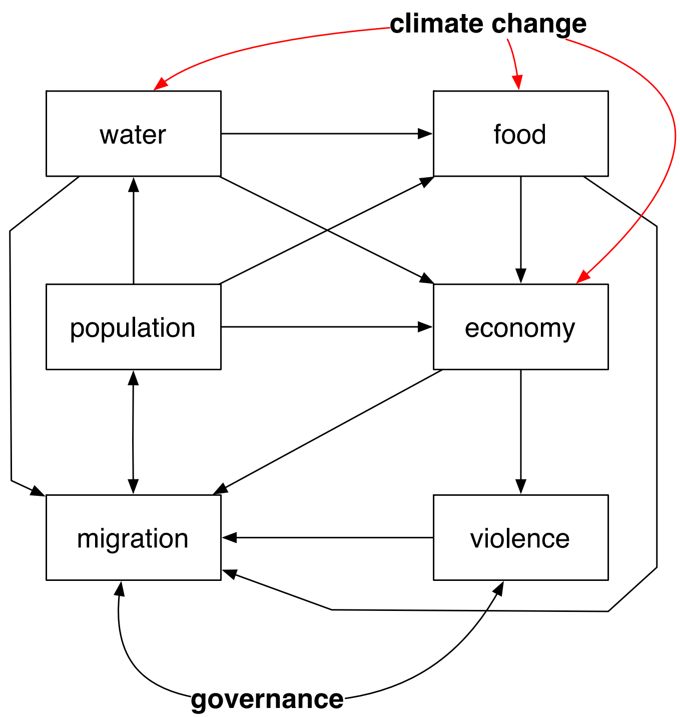


Figure 6.1: Overview of model sector structure

Population dynamics are simulated using growth rates calibrated to match historical data and projections (United Nations 2012). The Malian population is separated by gender (male and female), by age group (0-14, the potentially productive and fighting-aged workforce aged 15-64, and 65 and over), and by type of labor (skilled and common, with a static fraction for each determined exogenously) for the 15-64 age range. The stock and flow structure that determines population dynamics includes births, mortality for each age block, and aging flows.

The model simulates migration by determining the fraction of the Malian population that chooses to live in each simulated region. Potential locations are urban Mali, rural Mali, neighboring countries (including Burkina Faso, Cote d'Ivoire, Gabon, Gambia, Ghana, Guinea, Mauritania, Niger, Nigeria, and Senegal), the United States, and the rest of the world. Decisions about where to live are calculated using a cognitive formulation based on qualitative choice theory (McFaddin 1982). Utility functions for living in each region are calculated using equation 1.

$$\begin{aligned} MU_{i,l,g,v} = {} & \alpha_{i,l,g,v} + \beta_{1i,l,g,v} * \ln\left(\frac{WI_i}{WI_{oi}}\right) + \beta_{2i,l,g,v} * \ln\left(\frac{FA_i}{FA_{oi}}\right) + \beta_{3i,l,g,v} * \ln\left(\frac{GG_i}{GG_{oi}}\right) \\ & + \beta_{4i,l,g,v} * \ln\left(\frac{GI_i}{GI_{oi}}\right) + \beta_{5i,l,g,v} * \ln\left(\frac{DM_i}{DM_{oi}}\right) + \beta_{6i,l,g,v} * \ln(VI_i) \\ & + \beta_{7i,l,g,v} * \ln(II_i) + \beta_{8i,l,g,v} * \ln(FI_i) + \beta_{9i,l,g,v} * \ln\left(\frac{UER_{i,l}}{UER_{oi,l}}\right) \\ & + \beta_{10i,l,g,v} * \ln\left(\frac{POP_{i,x,l,g,m}}{POPx_{l,g,m}}\right) + \beta_{11i,l,g,v} * \ln(NDI_t) \end{aligned} \quad (1)$$

Where MU = migration utility
i = region
x = region of origin
l = labor type (skilled or common)
g = gender
v = age group
α = baseline utility
β_j = weight given to associated input
0 = initial value
WI = wage income
FA = food availability
GG = governance effectiveness
GI = infrastructure and services provided by government
DM = disease mortality
VI = violence incongruity
II = income incongruity
FI = food incongruity
UER = unemployment rate
POP = population
NDI = natural disaster index

All of the inputs to the utility functions are simulated dynamically within the model. The α factor is static but different for each region, labor type, gender, and age group. This represents any relatively steady factors not included explicitly as inputs, including culture. Incongruity factors are cognitive interpretations comparing current to baseline values. Weights for the utility functions are determined through a calibration process. The utility functions are then compared using a multinomial logit function (equation 2) that determines the fraction of Malians choosing to live in each region. This process is repeated at each time step to determine how the population in each region changes over time, simulating migration.

$$MP_{i,l,g,v} = Exp(MU_{i,l,g,v}) / \textstyle\sum_k Exp(MU_{k,l,g,v}) \quad (2)$$

Where MP = migration probability
k = the set of regions

Violence is calculated in a similar manner to migration. Decisions about whether to participate in violence are based on existing violence, governance effectiveness, food availability, unemployment, and wages.

The model simulates the economic situation in each region using a Cobb-Douglas formulation. We first calculate the potential gross regional product using labor supply, split into skilled and common labor, from the population sector of the model, capital, and technology.

$$PGRP_i = PGRP_0 * \left(\frac{LS_{i,s}}{LS_{0i,s}}\right)^{LsF_i} * \left(\frac{LS_{i,c}}{LS_{0i,c}}\right)^{LcF_i} * \left(\frac{C_i}{C_{0i}}\right)^{CF_i} * \left(\frac{TK_i}{TK_{0i}}\right)^{(LcF_i + LsF_i)} \quad (3)$$

Where PGRP = potential gross regional product
LS = labor supply
s = skilled labor
c = common labor
LsF = skilled labor fraction
LcF = common labor fraction
C = capital
CF = capital fraction
TK = technology

Capital is calculated based on depreciation and investment in capital, which is determined by the strength of the economy as well as governance effectiveness and infrastructure and services provided by the government. Technology is a calibration parameter, chosen to allow an initialization simulation of potential gross regional product to track data (World Bank 2014) and projections (IPCC 2000, scenario B2). Although somewhat dated, the IPCC 2000 projections provide a consistent and complete data set amenable to the purposes of this study. Simulation results are meant to illustrate the consequence of economic and societal interactions as a result of climate change. The IPCC 2000 data set used in the base case acts as a referent to establish a basis for comparison to alternative conditions. Realized gross regional product utilizes the potential gross regional product, and is altered by elasticity parameters for temperature, extreme events, water availability, food availability, governance effectiveness, and infrastructure and services provided by the government.

$$RGRP_i = PGRP_i * EC_i^{\alpha_i} * EG_i^{\beta_i} * EL_{s,i}^{\gamma_i} * EL_{c,i}^{\delta_i} * RA_i^{\epsilon_i} * GG_i^{\mu_i} * GI_i^{\sigma_i} \quad (4)$$

Where RGRP = realized gross regional product
EC = effective capital
EG = effective land
EL = effective labor
RA = resource availability
α, β, γ, δ, ε, μ, σ are sensitivity factors for the associated inputs

The realized gross domestic product for each region, along with population dynamics, helps to determine labor and wage dynamics. As the realized gross domestic product increases, employment increases. Employment is compared to labor supply to determine the unemployment rate. Wages are dependent on the fraction of skilled versus common labor (exogenous to the model), employment rates, and technology. Malians who have moved away send remittances back to the country, which adds to the total per capita income for those remaining in Mali. Generic resource availability is also calculated, based on extraction and use of resources and generation rates. Resource utilization is dependent on capital, technology, governance effectiveness, and infrastructure and services provided by the government.

6.4. Results

Initial results of the Malian migration model show a greater proportion of the population migrating outside of the country when temperatures rise. The base case simulation, in which temperatures stay stable throughout the time horizon and gross regional product tracks data, is shown in figures 6.2 and 6.3. Figure 6.2 shows the population of Malians in each of the five modeled regions: urban Mali, rural Mali, neighboring countries, the United States, and the rest of the world. The population rises fastest in urban Mali. Rural Mali and neighboring countries both see a substantial population increase, while the United States and the rest of the world remain relatively stable.

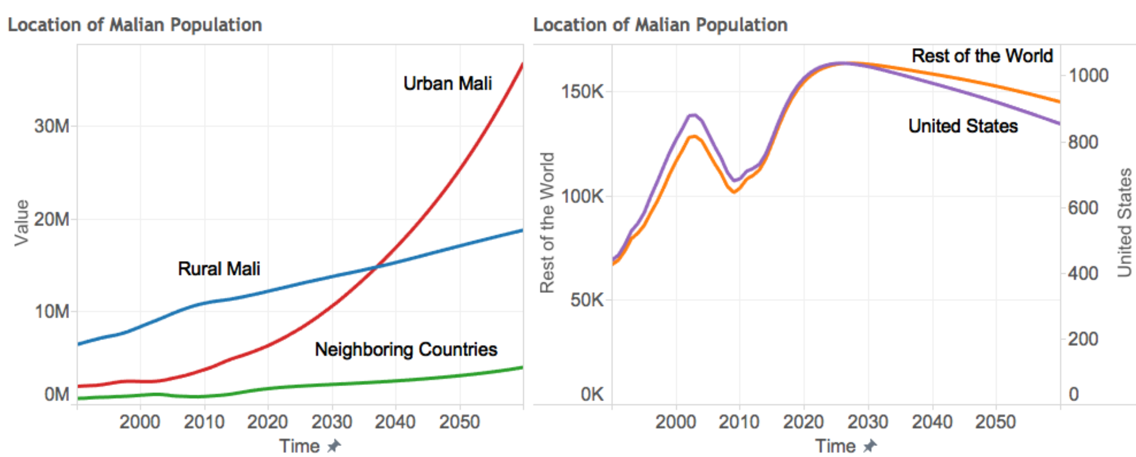


Figure 6.2: Location of Malian population in base case scenario

The most influential drivers of migration decisions in the model are violence, unemployment rates, income, and the effectiveness of governance in each region. In the base case, all regions are beneficially affected by projections of improved economic growth (figure 6.3). This analysis does not take a position on the validity of the base case projection. Because it provides for continuity and consistency between historical and future conditions, the difference between the climate change scenario and base case results can consistently portray the causes of variations in the results. Violence in all regions remains relatively stable in the base case scenario, with some fluctuation in response to wages, unemployment, and the effectiveness of government. Unemployment rates remain relatively stable, although they drop in urban Mali and neighboring countries where gross regional products rise substantially over the time horizon. Wage income in all areas increases throughout the time horizon. Governance effectiveness, an exogenous variable based on a corruption index (World Bank 2014), are the same over all scenarios. The gross regional product for Mali in the base case is calibrated to track projected data values, and drives all other economic variables including unemployment rates and wage income.

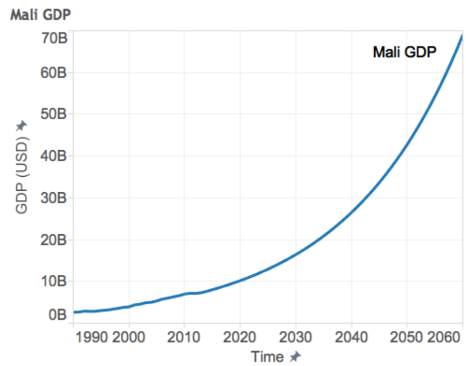


Figure 6.3: Malian GDP in base case scenario

The model can address the impact of climate-induced temperature, precipitation, and extreme weather on economic and demographic conditions. Although precipitation is typically used for highlighting the impact of climate change in agriculturally dependent countries, the scenario presented here considers the often-neglected impacts of temperature, focusing on its impact for labor effectiveness (Dunne 2013). The climate change scenario (figure 6.4) assumes a linear temperature increase from 2010 through 2100. The total temperature change for each region over this time horizon is 2.5°C for Mali and neighboring countries, 4°C for the United States, and 3°C for the rest of the world, due to relative changes in latitude. We assume that temperature influences the economy by changing effective capital, land, and labor. Sensitivities of each of these variables to temperature are variable by region, with the United States reacting the least to climate change and Mali reacting most heavily. Gross regional product is also strongly affected by migration patterns, with larger populations providing more labor with a potential boost to the economy. Mali's gross regional product is lower than in the base case by the end of the time horizon.

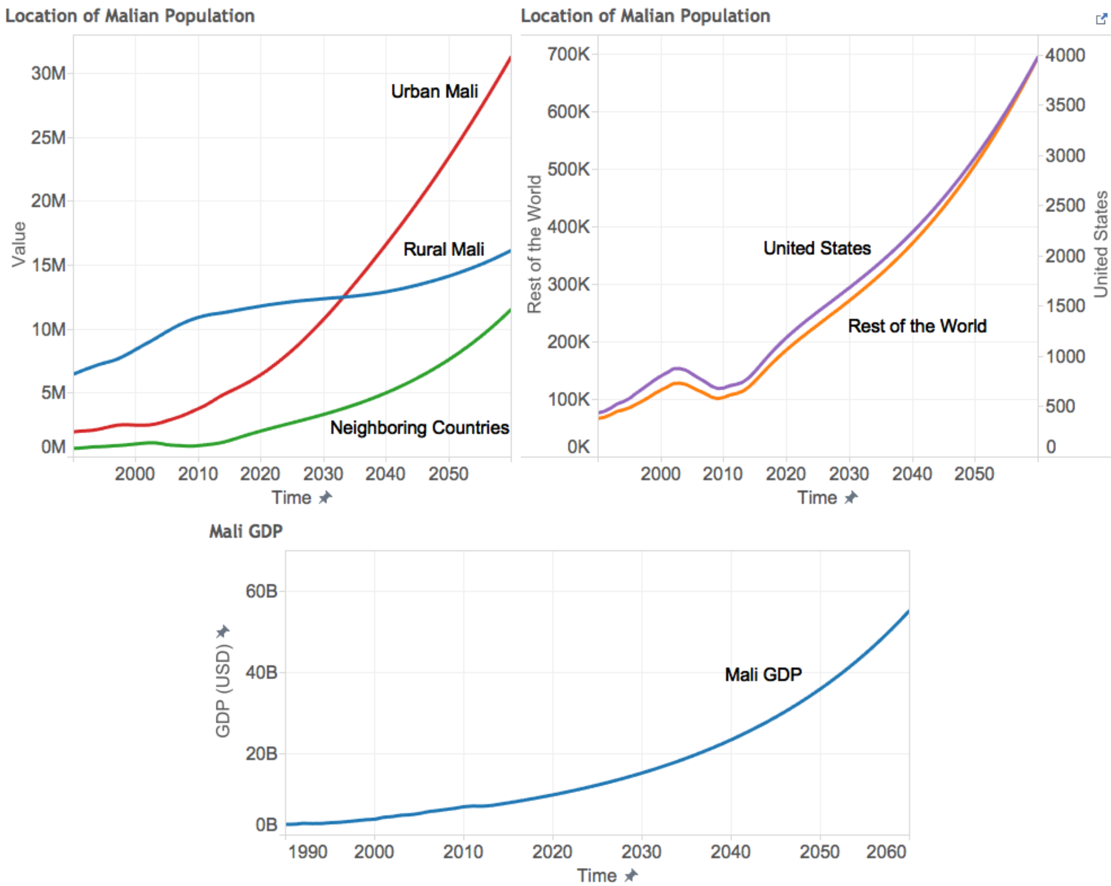


Figure 6.4: Climate change scenario

Wage income for neighboring countries, the United States, and the rest of the world is lower in the climate change case than in the base case due to the drop in GDP from climate impacts. Urban and rural Mali actually both see a slight increase in wage income in this scenario. As the population moves away from Mali, remittances increase, so even with a lower gross regional product Malians see higher effective income. Unemployment rates in all regions increases due to lower economic activity. Violence increases, both over the time horizon and in comparison to the base case, as unemployment increases and wages either decrease or do not increase enough to counteract unemployment rates.

While the total population of Malians is the same in the climate change case and the base case, urban and rural Malian populations are both lower in this scenario than in the base case. As the gross regional product in Mali drops and unemployment increases, the population moves to neighboring countries, the United States, and the rest of the world. The population in neighboring countries increases substantially, surpassing 10 million by the end of the time horizon. United States and rest of the world populations of Malians increase steadily, a distinct difference from the base case in which they remain relatively stable.

The migration model uses weights to define how various inputs affect utility functions that ultimately determine migration decisions. We calibrated the model to data and projections where possible, but these weights could be defined differently. To explore the range of potential outcomes, we conducted an uncertainty quantification analysis on weights that determine the effects of populations, effectiveness of governance, unemployment rates, wage income, and violence on migration decisions. Results of this analysis (figure 6.5) show that the basic patterns of population movement remain the same. The base case sees a wider spread of potential outcomes given the different cognitive weights. United States and rest of the world populations remain relatively steady. Populations in urban Mali and neighboring countries increase. The rural Malian population may exhibit a decrease in the second half of the time horizon given certain weight combinations, which would results in higher populations in the other regions. Patterns of population growth in the climate change case are more robust. Populations in urban Mali, neighboring countries, the United States, and the rest of the world increase in all simulations. The population of rural Mali tends to remain more even than in the base case.

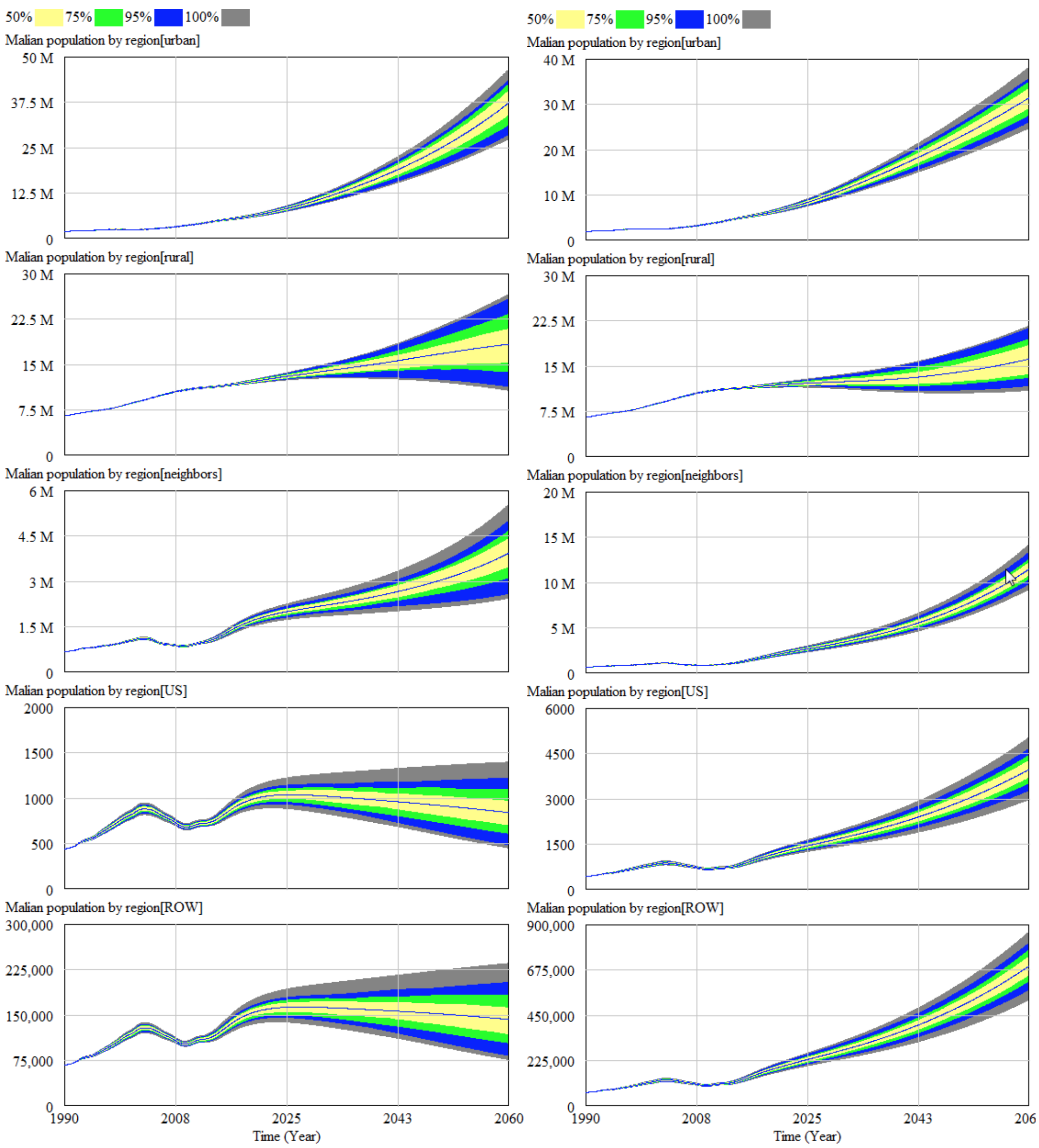


Figure 6.5: Sensitivity results for population of Malians in each region with variation in cognitive parameters: base case (left side) and climate change case (right side)

6.5. Policy Analysis

We used the model to explore five policy options for reducing migration from Mali in response to climate change. These policy options were increasing contraception availability (implemented by a reduction in the modeled birth rate), increases in governance effectiveness and in infrastructure and services provided by the government, and increasing foreign aid to either urban or rural areas in Mali (implemented through an increased potential gross regional product). Figure 6.6 shows the results of a series of simulations with climate change turned on and variation in the policy variables. The policy variable ranges were as shown in table 6.1. For this ensemble, three evenly spaced points (including the low and high values) were chosen for each variable, giving a total of 243 simulations. Since all of the policy variables were chosen to decrease the need for out migration, the output of this ensemble of runs tends to show decrease from the base case (the blue line in each chart in figure 6.6) for Malians living in neighboring countries, the US, and the rest of the world. Urban and rural populations in the simulations can be either higher or lower. This is because some policy variables, such as economic aid to rural areas, are meant to improve conditions in either the urban or rural areas, but not necessarily to both evenly. Thus, some simulations lead to a large increase in the urban population but a drop in the rural population, and vice versa.

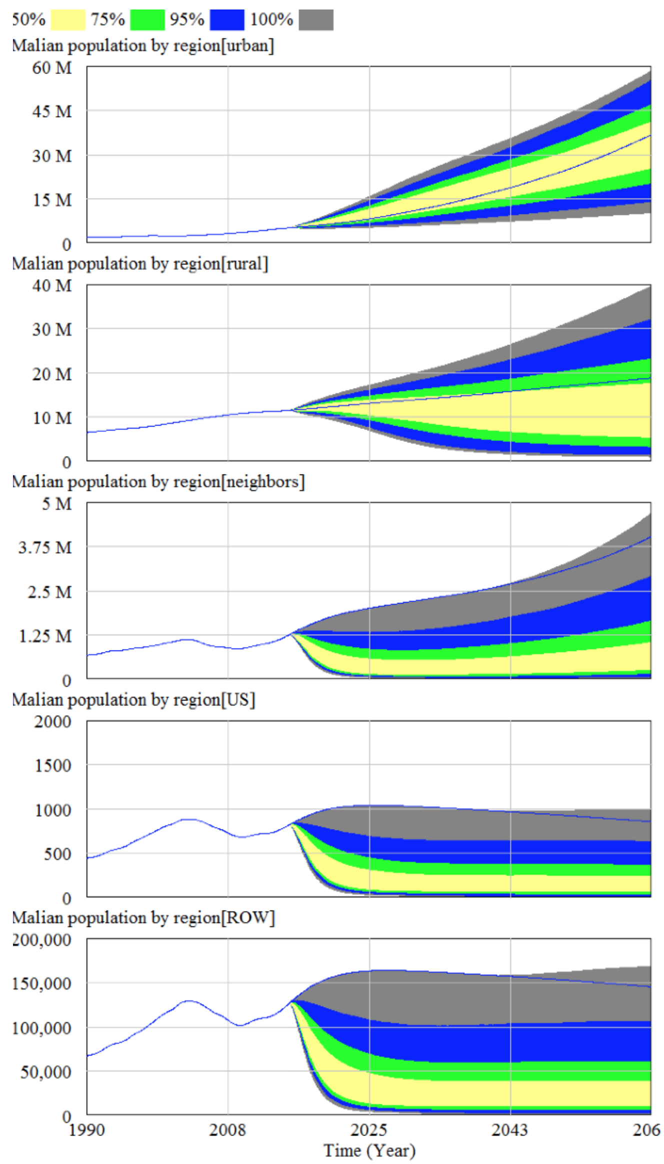


Figure 6.6: Output of ensemble of runs with climate change on and variation in policy variables

Policy	Variable	Low Value	High Value
contraception	percent reduction in birth rate	0	50
governance effectiveness	percent increase in governance effectiveness	0	100
infrastructure/ services	percent increase in infrastructure/ services provided by the government	0	100
aid to urban areas	RGRP increase from economic aid (urban)	0	100,000,000
aid to rural areas	RGRP increase from economic aid (rural)	0	100,000,000

Table 6.1: Policy variable ranges

Sensitivity analysis results in the form of correlation coefficients over time for the five policies are shown in Figure 6.7. The outputs of interest for this sensitivity analysis were the total population of Mali, the population of urban Mali, and the population of rural Mali. Urban and rural populations are most strongly affected by economic aid policies. Economic aid to urban areas will increase the population in those areas, but will also decrease the population in rural areas. Conversely, economic aid to rural areas will increase the rural population while decreasing the urban population, since the policy creates an incentive for people living in urban areas to move to rural areas. Contraception is a major driver of both sub-populations, especially in the later part of the simulation, when a lower birth rate over time has impacted the population more through exponential growth. Governance effectiveness affects the rural population toward the beginning of the time horizon, before other variables create a stronger pull. Sensitivity analysis results for the total population in Mali are substantially different than the results for the sub-populations. This is because the total population ignores movement between the urban and rural populations. For the total population analysis, contraception is the strongest driver. Governance effectiveness also strongly affects the results, particularly at the beginning of the time horizon.

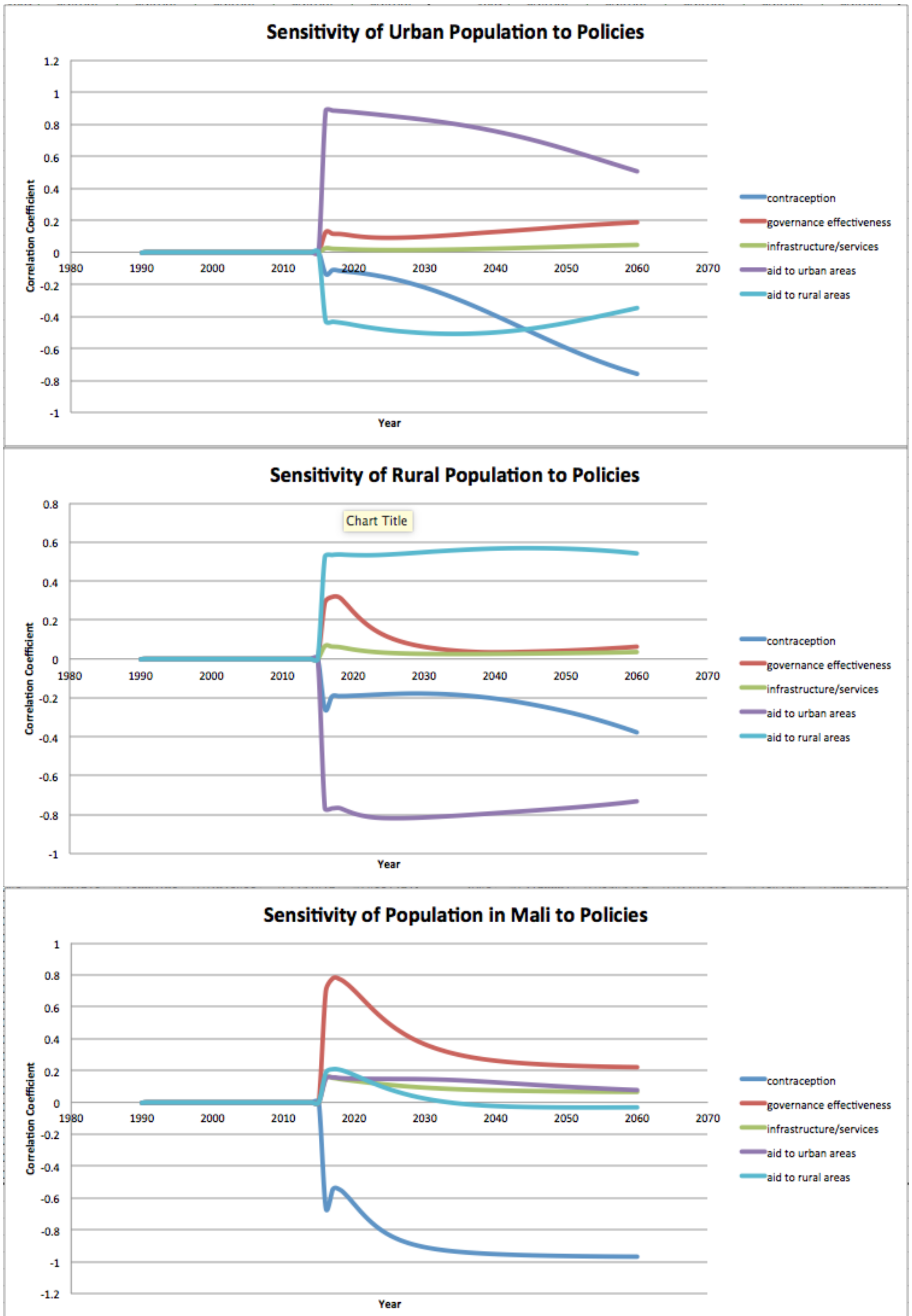
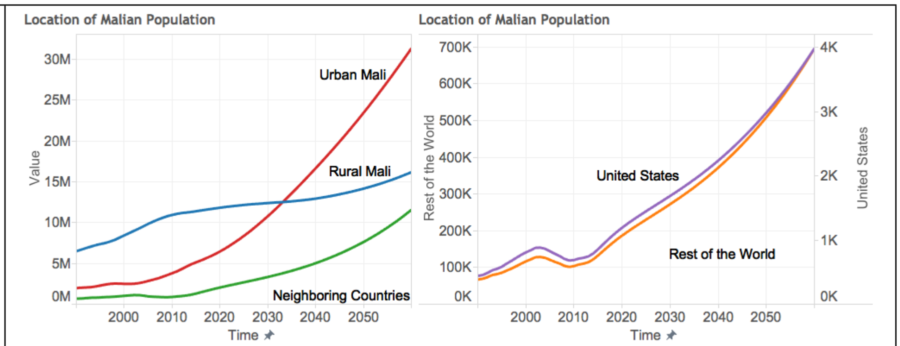


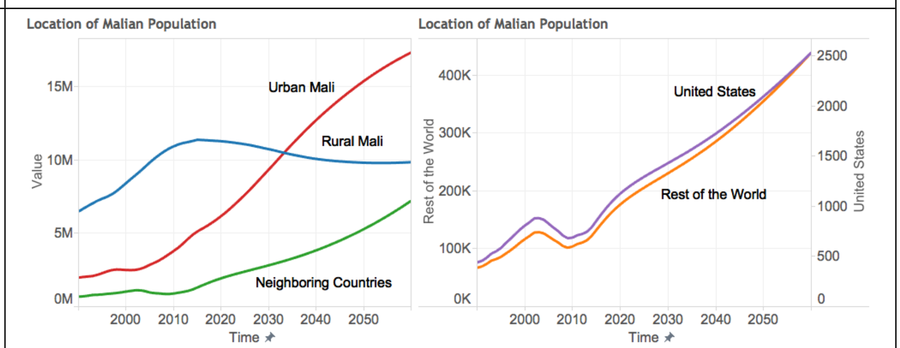
Figure 6.7: Correlation coefficients over time for policy options in relation to the urban population, rural population, and total population in Mali

Figure 6.8 shows results for the population of the five areas over time, for seven potential policy options. Only the contraception option results in all of the populations decreasing as compared to the no-policies case. As indicated in the sensitivity analysis, increasing governance effectiveness is the most effective at decreasing the number of Malians migrating out of the country. Economic aid to urban and/or rural areas tends to draw the population toward those areas. Higher amounts of economic aid were not considered in this analysis, but could be analyzed with the model, and would likely lead to stronger pulls for migration into the receiving areas.

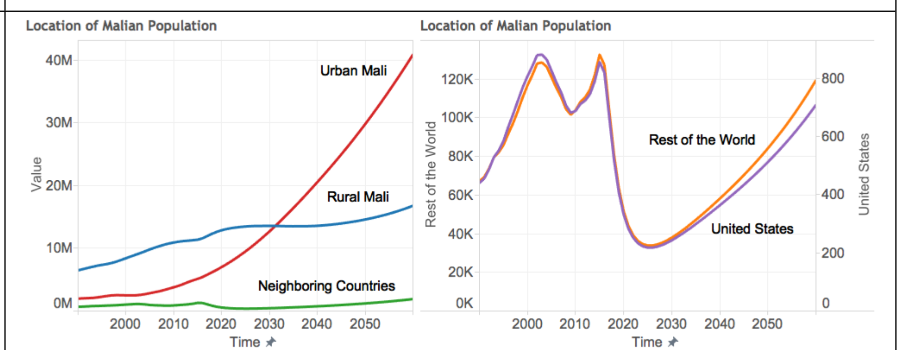
Climate change with no policies



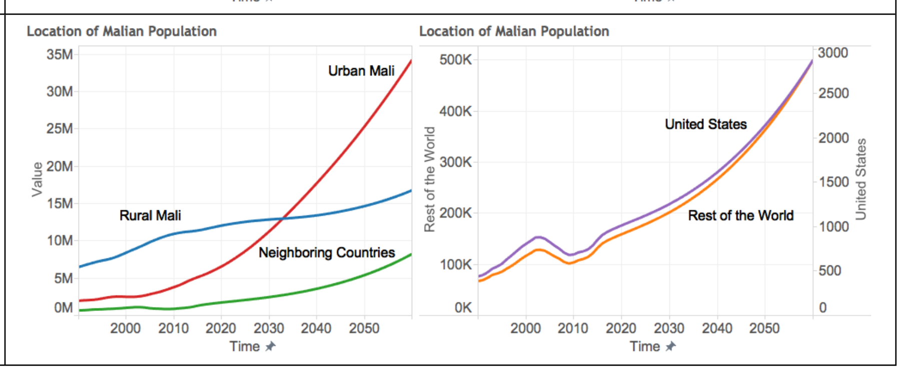
Birth rate reduced 50% (contraception)



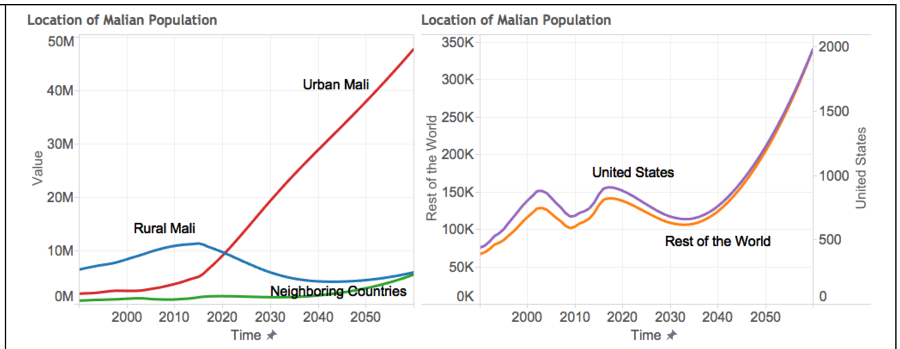
Governance effectiveness increased 100%



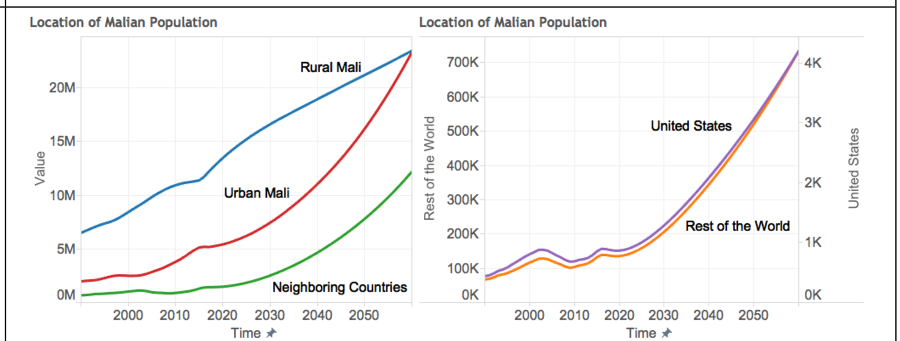
Government infrastructure/services increased 100%



$100,000,000 aid to urban areas



$100,000,000 aid to rural areas



$50,000,000 aid to urban areas and $50,000,000 aid to rural areas

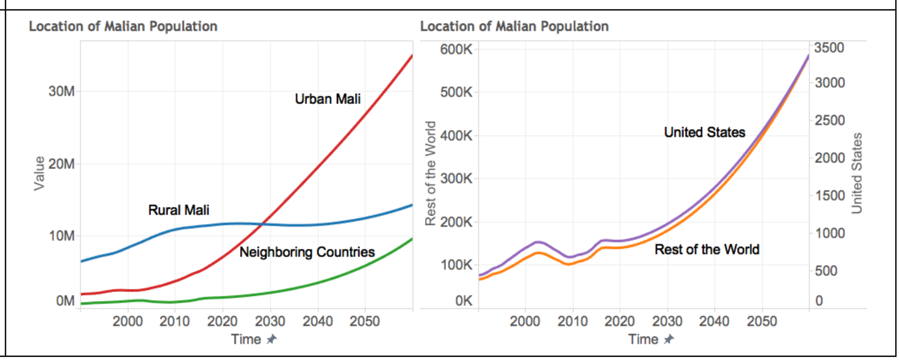


Figure 6.8: Results for population in each location given different policy options

6.6. Conclusions

Understanding likely migration patterns can improve national security and human aide capabilities. The system dynamics-based model described here couples migration behavior with the interacting dynamics of economy, labor, population, violence, governance, water, and food. We use qualitative choice theory to represent migration decisions, and economic theory to simulate major drivers of these decisions. This model shows a relatively robust pattern of migration of the Malian population in response to climate change. As temperatures increase, economic factors make migration from Mali to other locations more attractive. The population tends to move out of both urban and rural areas of Mali, and toward neighboring countries, the United States, and the rest of the world.

This regional model focused on Mali in western Africa has been adopted for the first test case. A variety of additions could improve the utility of this model. We have not yet modeled migration decisions of populations other than Mali. For example, a better understanding of population dynamics for neighboring countries by coupling detailed migration models of multiple countries or regions could improve our model of economic factors and thus migration decisions of Malians. We also plan to improve the food, water, and disease sectors of the model. Sensitivity analysis could help us to determine which cognitive weights are most impactful to the model, and further research could help us to improve the accuracy of those weights. This test case shows that simulation of this type is promising, and might provide useful insight into likely effects of climate change on migration.

6.7. Acknowledgements

Support for this project was made possible by a grant from Sandia National Laboratories Laboratory Directed Research and Development Program. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

6.8. References

Afifi, T. and Warner, K., 2008. *The Impact of Environmental Degradation on Migration Flows Across Countries, United Nations Working Paper No. 5/2008*, Bonn.

Azam, J.P., & F. Gubert. 2006. Migrants' Remittances and the Household in Africa: A Review of the Evidence. *Journal of African Economies,* Vol. 15, AERC Supplement 2: 426-462.

Black, R. 1998. *Refugees, Environment and Development*. London: Longman. Black, R. 2001.

Black, R., 2001. Environmental Refugees: Myth or Reality? New Issues in Refugee Research Working Paper 34. Geneva: UN High Commissioner for Refugees

Black, R.B., Adger, N., Arnell, N., Dercon, S., Geddes, A., Thomas, D., 2011. *Foresight: Migration and Global Environmental Change, Final Project Report*, The Government Office for Science, London.

Brown, O., 2008. *Migration and Climate Change, International Organization for Migration (IOM): Research Series No. 31*, IOM Geneva.

Chichilnisky, G. and DiMatteo, M, 1998. *Trade, Migration and Environment: A General Equilibrium Analysis*, G. Chichilnisky, G. Heal, and A Vercelli (eds). Kluwer Academic Publishers.

Christian Aid, 2007. *Human Tide: The real migration crisis,* Christian Aid Report May 2007. London.

Defence Science Board, 2011. *Trends and Implication of Climate Change for National and International Security*.

Davis, K.F., D'Odorico, P., Laio, F., Ridolfi, L., 2013. Global Spatio-Temporal Patterns in Human Migration: A Complex Network Perspective, *PLoS ONE*, Vol. 8, No. 1.

Dunne, J. P., Stouffer, R. J., and John, J. G., 2013. Reductions in labour capacity from heat stress under climate warming. *Nature Climate Change*. Vol. 3: 563-566.

El-Hinnawi, E., 1985. *Environmental Refugees*. United Nations Environment Programme, Nairobi.

Ghimire, D.J. and Mohai, P.L., 2005. Environmentalism and Contraceptive Use: How People in Less Developed Settings Approach Environmental Issues, *Population and Environment*, Vol. 27, No. 1: 29-61.

Gilbert, G., and McLeman, R., 2010. Household access to capital and its effects on drought adaptation and migration: A case study of rural Alberta in the 1930s. *Population and Environment*, Vol.32, No.1: 3-26.

Grote, U., Engel, S., Schraven, B., 2006. *Migration due to the Tsunami in Sri Lanka –Analyzing Vulnerability and Migration at the Household Level*. ZEF-Discussion Paper on Development Policy No. 106, Centre for Development Research (ZEF), Bonn, April 2006.

Homer-Dixon, Thomas F. 1991. On the Threshold: Environmental Changes as Causes of Acute Conflict. *International Security* 16: 76-116.

Intergovernmental Panel on Climate Change (IPCC). 2000. *Emissions Scenarios*. Ed. Nebojsa Nakicenovic and Rob Swart, Cambridge University Press, Cambridge, England. http://www.ipcc.ch/ipccreports/sres/emission/

Internal Displacement Monitoring Center and the Norwegian Refugee Council (IDMC), 2011. *Internal Displacement: Global Overview of Trends and Developments in 2010*.

Kniveton, D.R., Smith, C.D., Black, R., 2012. Emerging Migration Flows in a Changing Climate in Dryland Africa, Nature Climate Change Letters.

Koubi, V., Spilker, G., Schaffer, L. M. and Bernauer, T., 2012. Environmental Degradation and Migration. Available at SSRN: http://ssrn.com/abstract=2107133 or http://dx.doi.org/10.2139/ssrn.2107133

Massey, D., Axinn, W., and Ghimire, D., 2007. *Environmental Change and Out-Migration: Evidence from Nepal*, Report 07-615, Population Studies Center, University of Michigan, 28p.

Massey, D.S. Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., Taylor, J.E., 1993. Theories of Migration: A Review and Appraisal, *Populating and Development Review*, Vol. 19, No. 3: 431-466.

McFadden, D., 1982. *Qualitative Response Models*. In Advances in Econometrics, Ed. Werner Hildenbrand, Cambridge University Press, New York.

McLeman, R., and Smit, B., 2006. Migration as an Adaptation to Climate Change. *Climatic Change*, Vol. 76, No.1-2: 31-53.

Mena, C.F., Walsh, S.J., Frizzelle, B.G., Yao, X.Z., and Malanson, G.P., 2011. Land use change on household farms in the Ecuadorian Amazon: Design and implementation of an agent-based model, Applied Geography, 31(1): 210-222.

Myers, N., 2002. Environmental Refugees: A Growing Phenomenon of the 21st Century, *Philosophical Transactions of The Royal Society B*. Vol. 357: 609-613.

National Intelligence Council, 2012. *Global Trends 2030: Alternative Worlds*. NIC 2012-001.

Perch-Nielsen, S.L., 2004. *Understanding the effect of climate change on human migration: The contribution of mathematical and conceptual models, diploma thesis*. Swiss Federal Institute of Technology, Zurich.

Reuveny, R. and Moore, W.H., 2009. Does environmental degradation influence migration? Emigration to developed countries in the last 1980 and 1990s, *Social Science Quarterly*, 90(3): 461-479.

Salehyan, I., and Gleditsch, K.S., Refugees and the Spread of Civil War, 2006. *International Organization,* Vol. 60: 335–366.

Siyaranamual, M.D., 2009. *Constructing the Model of Environmental Migration*, Working Paper in Economic and Development Studies No. 200901, Padjadjaran University, 25p.

United Nations Framework Convention on Climate Change, 2013. *Nairobi work programme on impacts, vulnerability and adaptation to climate change*, Decision 17/CP.19.

United Nations, 2009a. *International Migrant Stock: The 2008 Revision*. United Nations Population Division (http://esa.un.org)

United Nations. 2009b. *Monitoring Global Population Trends*. United Nations Population Division (http://www.un.org/en/development/desa/population/publications/urbanization/urban-rural.shtml)

United Nations, 2012. *World Population Prospects: The 2012 Revision*. http://esa.un.org/unpd/wpp/index.htm.

Ventana Systems, 2013. *Vensim DSS for Windows Version 6.3 Double Precision.* Ventana Systems, Inc.

Wood, W.B., 2001. *Ecomigration: Linkages Between Environmental Change and Migration*. Pp. 42-61 in Aristide R. Zolberg and Peter Benda, eds., Global Migrants, Global Refugees. New York and Oxford: Berghahn.

World Bank. 2014. *World Development Indicators*. http://data.worldbank.org/data-catalog/world-development-indicators. Accessed November 13, 2014.

APPENDIX A: STATE SPACE MODEL SPECIFICATION WITH ERRORS REPRESENTED AS STATE VARIABLE WITH AR(1) PROCESS

Descriptive Statistics

We first plot some of the figures of the relationship between dependent variable and independent variables to demonstrate how the log transformation of the Partial Adjustment Model (1) achieved Gaussian series.

Figure A1: Relationship between Log Income and Log Outmigration Flows: 1970-2010

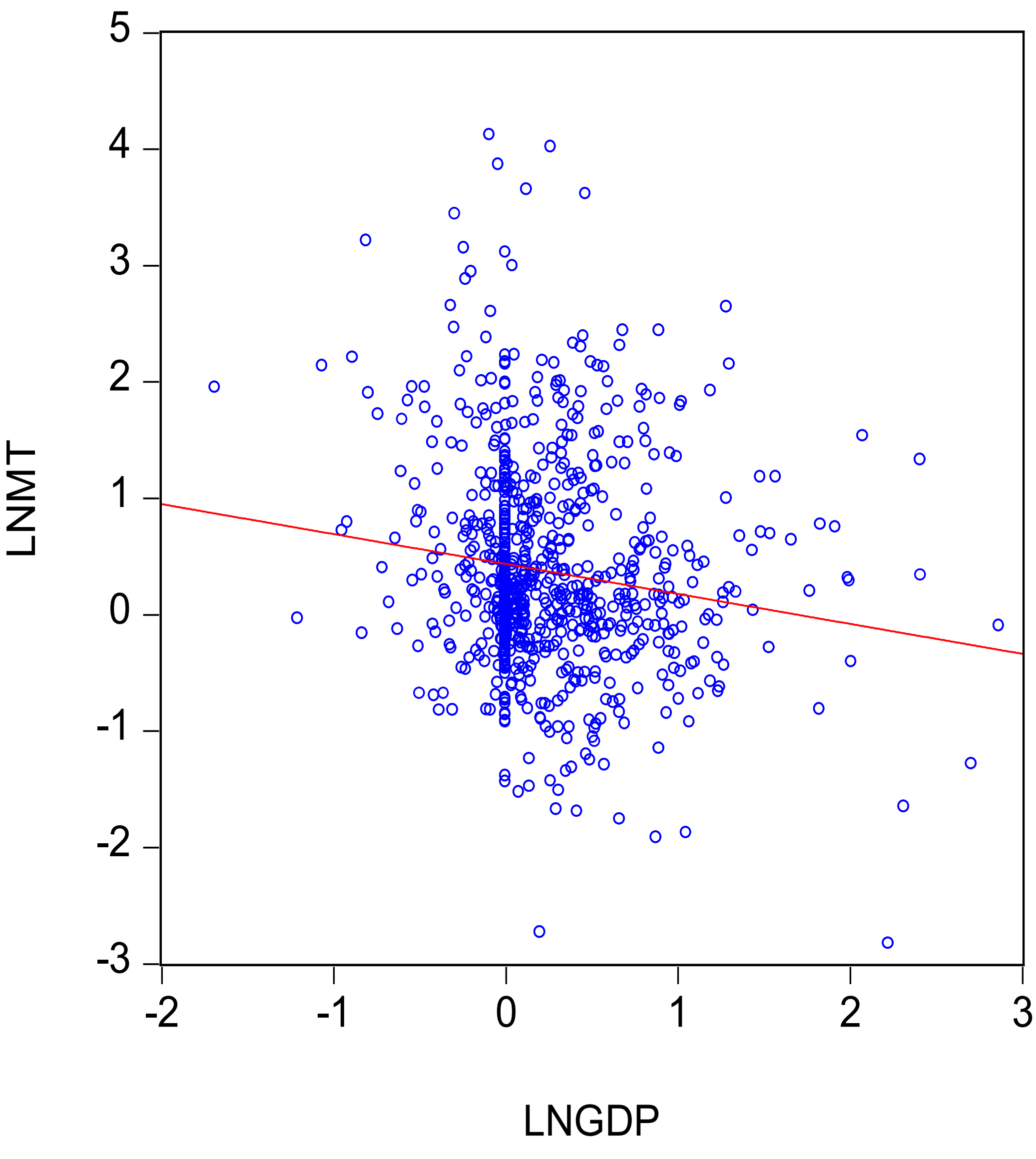


Figure A2: Relationship between Log of Cereal Production and Log Outmigration Flows: 1970-2010

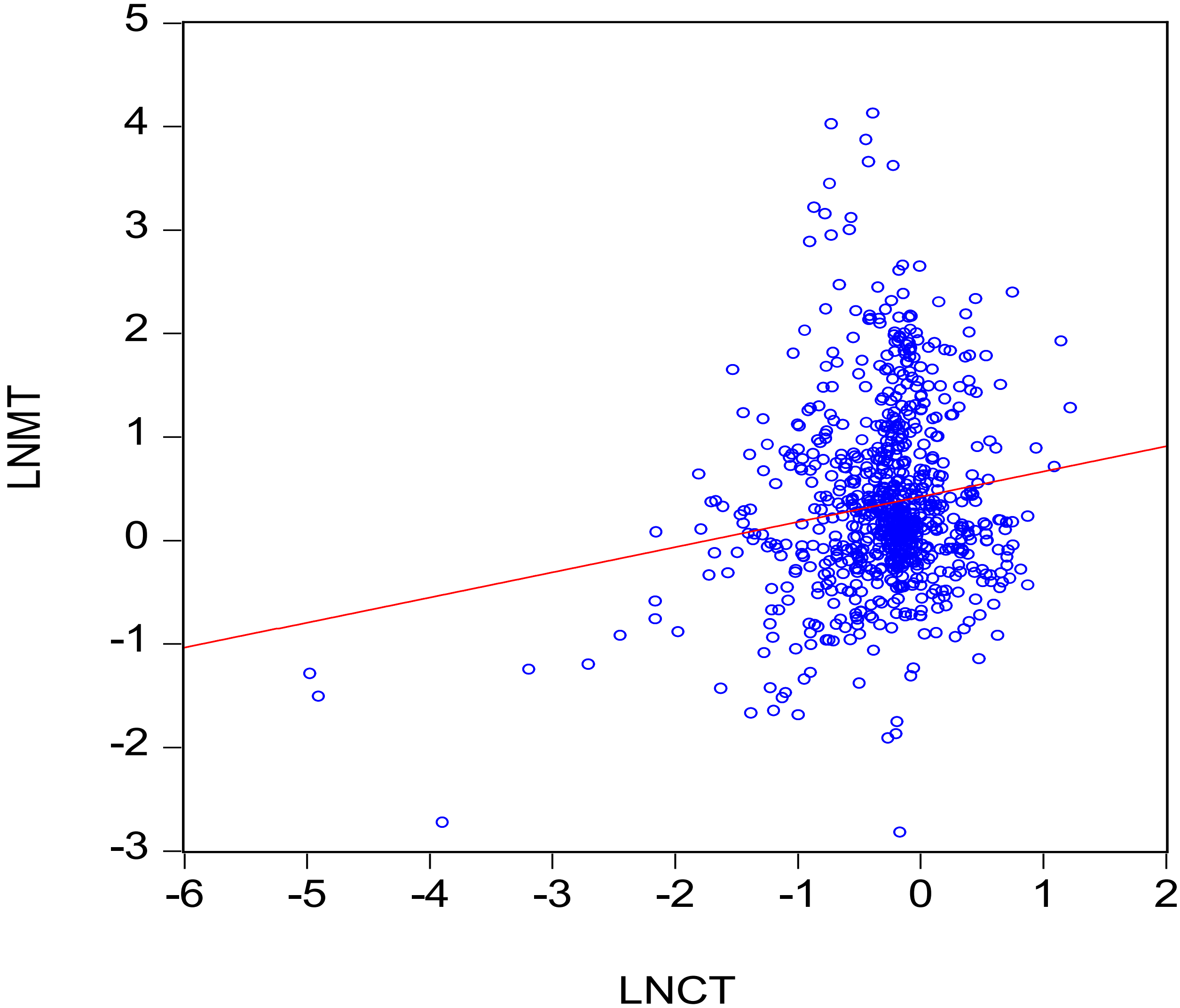


Figure A3: Relationship between Lagged Outmigration flows on Current Outmigration Flows: 1970-2010

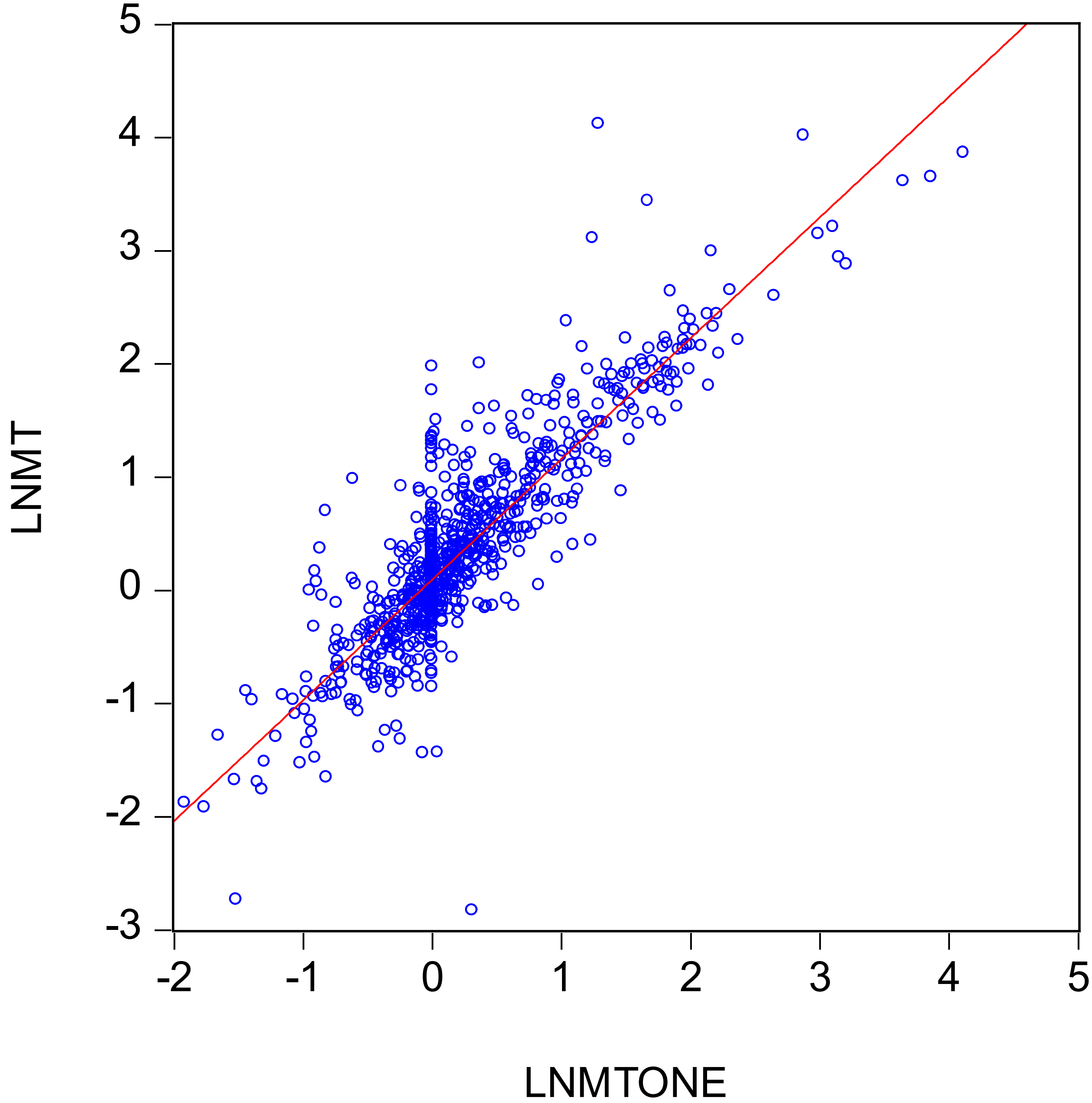
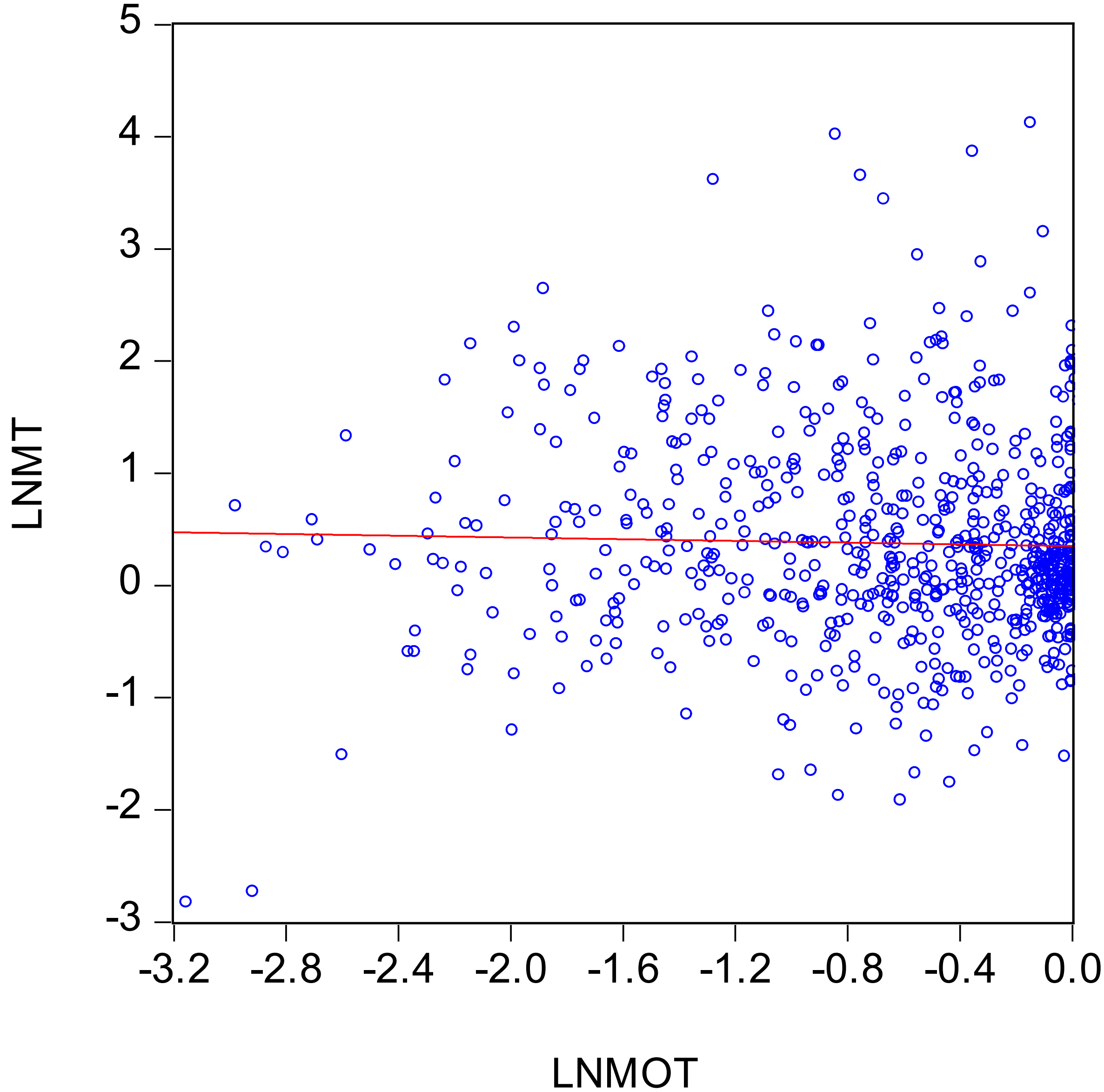


Figure A4: Relationship between Log of Infant Mortality Rate and Log Outmigration Flows: 1970-2010



All Nations Regression Analysis Results

As noted in the paper, multicollinearity between GDP and Adjusted Savings; GDP and Telephone Lines; Infant Mortality and Adjusted Savings rate; and, Infant Mortality and Telephone Lines were detected in the analysis. Thus, four sets of regression analysis were conducted, including GDP-Adjusted Savings combination, GDP-Telephone Lines combination, Infant Mortality-Adjusted Savings combination, and Infant Mortality-Telephone Lines combination. These results are reported in tables A.1 to A.4.

Table A.1: Regression Results for All Countries with GDP & Adjusted Savings Combinations — (Dependent Variable: Log of Emigration Flows for the Period 1970- 2010)

Variables	Common Coefficient I	Cross-section Specific Fixed Effects II	Panel GLS III	Panel GLS with Cross-section fixed effects IV
Constant	0.16*** (0.03)	0.142*** (0.037)	0.143*** (0.036)	0.13*** (0.036)
Log of Battle Deaths	0.012 (0.01)	0.014 (0.009)	0.01 (0.01)	0.012 (0.01)
Log of Cereal Production	0.104*** (0.024)	0.092*** (0.024)	0.089*** (0.024)	0.077*** (0.024)
Log of Natural Disasters	0.001 (0.008)	0.002 (0.008)	0.001 (0.007)	0.002 (0.007)
Log of GDP index	-0.137*** (0.05)	-0.133*** (0.05)	-0.111*** (0.04)	-0.107*** (0.04)
Log of Lagged Outmigration	1.06*** (0.019)	1.07*** (0.019)	1.054*** (0.02)	1.06*** (0.02)
Log of Adjusted Savings	0.015 (0.012)	0.018 (0.013)	0.01 (0.012)	0.012 (0.012)
Log of Polity Index	-0.022 (0.051)	-0.004 (0.052)	-0.007 (0.047)	0.009 (0.047)
Adjusted R^2	0.765	0.764	0.793	0.791
N	726	726	726	726
DW Statistic	1.95	1.96	1.957	1.96

Notes: ***, **, and * denotes significance at 1%, 5%, and 10% respectively.
The terms in the brackets denotes standard errors.

Table A.2: Regression Results for All Countries with GDP & Telephone Lines Combinations — (Dependent Variable: Log Emigration Flows for the Period 1970- 2010)

Variables	Common Coefficient I	Cross-section Specific Fixed Effects II	Panel GLS III	Panel GLS with Cross-section fixed effects IV
Constant	0.166*** (0.038)	0.193*** (0.043)	0.157*** (0.038)	0.186*** (0.045)
Log of Battle Deaths	0.005 (0.009)	0.009 (0.008)	0.006 (0.009)	0.009 (0.009)
Log of Cereal Production	0.084*** (0.029)	0.093*** (0.028)	0.083*** (0.026)	0.09*** (0.02)
Log of Natural Disasters	-0.001 (0.007)	-0.0012 (0.0078)	-0.001 (0.006)	-0.002 (0.007)
Log of GDP Index	-0.081** (0.037)	-0.087** (0.037)	-0.081** (0.034)	-0.085** (0.034)
Log of Lagged Outmigration	1.059*** (0.018)	1.054*** (0.02)	1.053*** (0.019)	1.048*** (0.02)
Log of Polity Index	0.02 (0.05)	0.009 (0.05)	0.029 (0.046)	0.019 (0.046)
Log of Telephone lines per capita	-0.028* (0.015)	-0.039** (0.017)	-0.027* (0.014)	-0.036** (0.016)
Adjusted R^2	0.773	0.773	0.793	0.793
N	782	782	782	782
DW Statistic	1.946	1.953	1.96	1.965

Notes: ***, **, and * denotes significance at 1%, 5%, and 10% respectively.
The terms in the brackets denotes standard errors.

Table A.3: Regression Results for All Countries with Infant Mortality and Adjusted Savings Combinations — (Dependent Variable: Log of Emigration Flows for the Period 1970- 2010)

Variables	Common Coefficient I	Cross-section Specific Fixed Effects II	Panel GLS III	Panel GLS with Cross-section fixed effects IV
Constant	0.154*** (0.037)	0.154*** (0.038)	0.139*** (0.036)	0.137*** (0.037)
Log of Battle Deaths	0.011 (0.01)	0.013 (0.01)	0.009 (0.01)	0.01 (0.01)
Log of Cereal Production	0.078*** (0.024)	0.078*** (0.023)	0.069*** (0.025)	0.066*** (0.024)
Log of Natural Disasters	4.5E-05 (0.007)	0.0001 (0.0086)	0.0004 (0.007)	0.001 (0.007)
Log of Lagged Outmigration	1.081*** (0.018)	1.081*** (0.018)	1.06*** (0.02)	1.07*** (0.02)
Log of Adjusted Savings	0.012 (0.012)	0.011 (0.012)	0.006 (0.012)	0.005 (0.012)
Log of Polity Index	-0.002 (0.056)	-0.001 (0.054)	0.006 (0.051)	0.01 (0.04)
Log of Infant Mortality	0.088*** (0.032)	0.087** (0.039)	0.07*** (0.027)	0.064** (0.032)
Adjusted R^2	0.765	0.764	0.791	0.79
N	726	726	726	726
DW Statistic	1.942	1.944	1.949	1.951

Notes: ***, **, and * denotes significance at 1%, 5%, and 10% respectively.
The terms in the brackets denotes standard errors.

Table A.4: Regression Results for All Countries with Infant Mortality Rate and Telephones Lines Combinations — (Dependent Variable: Log of Emigration Flows for the Period 1970-2010)

Variables	Common Coefficient I	Cross-section Specific Fixed Effects II	Panel GLS III	Panel GLS with Cross-section fixed effects IV
Constant	0.161*** (0.037)	0.208*** (0.046)	0.154*** (0.037)	0.198*** (0.048)
Log of Battle Deaths	0.003 (0.009)	0.008 (0.009)	0.004 (0.009)	0.008 (0.009)
Log of Cereal Production	0.07*** (0.02)	0.086*** (0.024)	0.069*** (0.025)	0.082*** (0.026)
Log of Natural Disasters	-0.0001 (0.006)	-0.002 (0.007)	-0.001 (0.006)	-0.003 (0.007)
Log of Lagged Outmigration	1.07*** (0.018)	1.06*** (0.018)	1.06*** (0.018)	1.056*** (0.019)
Log of Polity Index	0.03 (0.05)	0.013 (0.05)	0.036 (0.047)	0.02 (0.04)
Log of Telephone lines per capita	-0.027* (0.015)	-0.042** (0.018)	-0.03** (0.015)	-0.042** (0.017)
Log of Infant Mortality	0.055** (0.027)	0.076** (0.033)	0.045* (0.024)	0.061** (0.029)
Adjusted R^2	0.77	0.771	0.79	0.79
N	789	789	789	789
DW Statistic	1.962	1.967	1.977	1.98

Notes: ***, **, and * denotes significance at 1%, 5%, and 10% respectively.
The terms in the brackets denotes standard errors.

Developing Nations Regression Analysis Results

To maintain consistency with the all nations' case, four sets of regression analysis were conducted, including GDP-Adjusted Savings combination, GDP-Telephone Lines combination, Infant Mortality-Adjusted Savings combination, and Infant Mortality-Telephone Lines combination. These results are reported in Tables A.5 to A.8.

Table A.5: Regression Results for Developing Countries with GDP and Adjusted Savings Combinations — (Dependent Variable: Log of Emigration Flows for the Period 1970- 2010)

Variables	Common Coefficient I	Cross-section Specific Fixed Effects II	Panel GLS III	Panel GLS with Cross-section fixed effects IV
Constant	0.175 (0.043)	0.148*** (0.042)	0.163*** (0.042)	0.149*** (0.041)
Log of Battle Deaths	0.004 (0.011)	0.006 (0.01)	0.004 (0.01)	0.005 (0.01)
Log of Cereal Production	0.124*** (0.043)	0.104** (0.044)	0.118*** (0.045)	0.101** (0.046)
Log of Natural Disasters	0.002 (0.008)	0.003 (0.008)	0.001 (0.008)	0.003 (0.008)
Log of GDP Index	-0.065 (0.04)	-0.062 (0.041)	-0.057 (0.041)	-0.056 (0.043)
Log of Lagged Outmigration	1.048*** (0.025)	1.055*** (0.025)	1.038*** (0.025)	1.043*** (0.025)
Log of Adjusted Savings	0.007 (0.014)	0.008 (0.015)	0.007 (0.014)	0.007 (0.015)
Log of Polity Index	0.019 (0.065)	0.051 (0.066)	0.027 (0.064)	0.056 (0.064)
Adjusted R^2	0.772	0.773	0.808	0.809
N	570	570	570	570
DW Statistic	1.958	1.976	1.93	1.944

Notes: ***, **, and * denotes significance at 1%, 5%, and 10% respectively.
The terms in the brackets denotes standard errors.

Table A.6: Regression Results for Developing Countries with GDP and Telephone Lines Combinations — (Dependent Variable: Log of Outward Migration Flows for the Period 1970- 2010)

Variables	Common Coefficient I	Cross-section Specific Fixed Effects II	Panel GLS III	Panel GLS with Cross-section fixed effects IV
Constant	0.179*** (0.042)	0.206*** (0.047)	0.174*** (0.041)	0.21*** (0.049)
Log of Battle Deaths	0.002 (0.012)	0.006 (0.012)	0.003 (0.011)	0.006 (0.011)
Log of Cereal Production	0.093** (0.047)	0.105** (0.046)	0.097** (0.048)	0.111** (0.049)
Log of Natural Disasters	-0.001 (0.007)	-0.002 (0.007)	-0.003 (0.007)	-0.003 (0.007)
Log of GDP Index	-0.028 (0.036)	-0.036 (0.038)	-0.035 (0.037)	-0.041 (0.038)
Log of Lagged Outmigration	1.042*** (0.023)	1.035*** (0.025)	1.032*** (0.022)	1.025*** (0.024)
Log of Polity Index	0.08 (0.062)	0.074 (0.063)	0.08 (0.06)	0.076 (0.06)
Log of Telephone lines per capita	-0.033** (0.015)	-0.045** (0.02)	-0.028* (0.015)	-0.042** (0.02)
Adjusted R^2	0.782	0.783	0.809	0.81
N	626	626	626	626
DW Statistic	1.948	1.96	1.956	1.97

Notes: ***, **, and * denotes significance at 1%, 5%, and 10% respectively.
The terms in the brackets denotes standard errors.

Table A.7: Regression Results for Developing Countries with Infant Mortality and Adjusted Savings Combinations — (Dependent Variable: Log of Emigration Flows for the Period 1970- 2010)

Variables	Common Coefficient I	Cross-section Specific Fixed Effects II	Panel GLS III	Panel GLS with Cross-section fixed effects IV
Constant	0.174*** (0.044)	0.149*** (0.046)	0.164*** (0.043)	0.147*** (0.044)
Log of Battle Deaths	0.004 (0.011)	0.006 (0.01)	0.004 (0.01)	0.005 (0.01)
Log of Cereal Production	0.109*** (0.04)	0.094** (0.039)	0.106*** (0.041)	0.091** (0.041)
Log of Natural Disasters	0.001 (0.008)	0.002 (0.008)	0.0004 (0.007)	0.002 (0.008)
Log of Lagged Outmigration	1.057*** (0.026)	1.061*** (0.025)	1.045*** (0.026)	1.048*** (0.026)
Log of Adjusted Savings	0.007 (0.014)	0.003 (0.015)	0.004 (0.014)	0.001 (0.014)
Log of Polity Index	0.02 (0.06)	0.05 (0.067)	0.032 (0.065)	0.056 (0.065)
Log of Infant Mortality	0.047 (0.029)	0.027 (0.036)	0.034 (0.029)	0.016 (0.035)
Adjusted R^2	0.772	0.772	0.808	0.808
N	570	570	570	570
DW Statistic	1.958	1.974	1.93	1.947

Notes: ***, **, and * denotes significance at 1%, 5%, and 10% respectively.
The terms in the brackets denotes standard errors.

Table A.8: Regression Results for Developing Countries with Infant Mortality and Telephone Lines Combinations — (Dependent Variable: Log of Emigration Flows for the Period 1970- 2010)

Variables	Common Coefficient I	Cross-section Specific Fixed Effects II	Panel GLS III	Panel GLS with Cross-section fixed effects IV
Constant	0.169*** (0.04)	0.194*** (0.049)	0.16*** (0.04)	0.194*** (0.053)
Log of Battle Deaths	0.001 (0.012)	0.004 (0.012)	0.001 (0.011)	0.005 (0.011)
Log of Cereal Production	0.09** (0.044)	0.10** (0.044)	0.092** (0.046)	0.103** (0.047)
Log of Natural Disasters	0.001 (0.007)	0.0003 (0.007)	-2.45E-05 (0.006)	-0.0005 (0.007)
Log of Lagged Outmigration	1.042*** (0.023)	1.038*** (0.025)	1.034*** (0.023)	1.029*** (0.024)
Log of Polity Index	0.086 (0.061)	0.082 (0.062)	0.095 (0.06)	0.088 (0.06)
Log of Telephone lines per capita	-0.034** (0.017)	-0.044** (0.019)	-0.032* (0.017)	-0.042** (0.019)
Log of Infant Mortality	0.01 (0.03)	0.022 (0.034)	0.006 (0.03)	0.018 (0.034)
Adjusted R^2	0.779	0.78	0.805	0.809
N	635	635	635	635
DW Statistic	1.972	1.983	1.984	1.994

Notes: ***, **, and * denotes significance at 1%, 5%, and 10% respectively.
The terms in the brackets denotes standard errors.

Asian Nations Regression Analysis Results

To maintain consistency with the all nations' case, four sets of regression analysis were conducted, including GDP-Adjusted Savings combination, GDP-Telephone Lines combination, Infant Mortality-Adjusted Savings combination, and Infant Mortality-Telephone Lines combination. These results are reported in Tables A.9 to A.12.

Table A.9: Regression Results for the Asian Sample with GDP and Adjusted Savings Combinations — (Dependent Variable: Log of Emigration Flows for the Period 1970- 2010)

Variables	Common Coefficient I	Cross-section Specific Fixed Effects II	Panel GLS III	Panel GLS with Cross-section fixed effects IV
Constant	0.23* (0.11)	0.184 (0.111)	0.224** (0.106)	0.229** (0.103)
Log of Battle Deaths	0.006 (0.007)	0.013 (0.009)	0.012 (0.008)	0.018* (0.01)
Log of Cereal Production	0.022 (0.025)	0.026 (0.023)	0.034 (0.039)	0.035 (0.036)
Log of Natural Disasters	-0.006 (0.023)	0.001 (0.025)	-0.012 (0.018)	-0.007 (0.021)
Log of GDP Index	0.027 (0.104)	0.028 (0.112)	0.019 (0.088)	0.015 (0.097)
Log of Lagged Outmigration	1.039*** (0.056)	1.068*** (0.053)	1.016*** (0.053)	1.032*** (0.054)
Log of Adjusted Savings	-0.048 (0.036)	-0.033 (0.035)	-0.027 (0.026)	-0.018 (0.028)
Log of Polity Index	0.129 (0.209)	0.163 (0.219)	0.104 (0.194)	0.135 (0.209)
Adjusted R^2	0.77	0.773	0.852	0.852
N	149	149	149	149
DW Statistic	1.85	1.91	1.858	1.90

Notes: ***, **, and * denotes significance at 1%, 5%, and 10% respectively.
The terms in the brackets denotes standard errors.

Table A.10: Regression Results for the Asian Sample with GDP and Telephone Lines Combinations — (Dependent Variable: Log of Emigration Flows for the Period 1970- 2010)

Variables	Common Coefficient I	Cross-section Specific Fixed Effects II	Panel GLS III	Panel GLS with Cross-section fixed effects IV
Constant	0.255** (0.119)	0.149 (0.11)	0.238** (0.105)	0.196* (0.104)
Log of Battle Deaths	0.006 (0.007)	0.013 (0.009)	0.013* (0.007)	0.018* (0.01)
Log of Cereal Production	0.014 (0.021)	0.021 (0.019)	0.025 (0.035)	0.026 (0.032)
Log of Natural Disasters	-0.007 (0.021)	-0.0009 (0.022)	-0.012 (0.075)	-0.007 (0.019)
Log of GDP Index	-0.011 (0.082)	-0.055 (0.092)	-0.012 (0.075)	-0.039 (0.082)
Log of Lagged Outmigration	1.018*** (0.05)	1.045*** (0.047)	0.994*** (0.045)	1.012*** (0.047)
Log of Polity Index	0.131 (0.198)	0.186 (0.212)	0.125 (0.181)	0.171 (0.2)
Log of Telephone lines per capita	-0.042 (0.027)	0.015 (0.032)	-0.026 (0.181)	0.009 (0.031)
Adjusted R^2	0.775	0.778	0.853	0.852
N	163	163	163	163
DW Statistic	1.73	1.79	1.791	1.84

Notes: ***, **, and * denotes significance at 1%, 5%, and 10% respectively.
The terms in the brackets denotes standard errors.

Table A.11: Regression Results for the Asian Sample with Infant Mortality and Adjusted Savings Combinations — (Dependent Variable: Log of Emigration Flows for the Period 1970- 2010)

Variables	Common Coefficient I	Cross-section Specific Fixed Effects II	Panel GLS III	Panel GLS with Cross-section fixed effects IV
Constant	0.226** (0.11)	0.187* (0.105)	0.214** (0.101)	0.228** (0.094)
Log of Battle Deaths	0.005 (0.008)	0.012 (0.01)	0.011 (0.008)	0.017* (0.01)
Log of Cereal Production	0.003 (0.02)	0.02 (0.025)	0.012 (0.034)	0.024 (0.038)
Log of Natural Disasters	-0.0008 (0.019)	0.003 (0.02)	-0.007 (0.014)	-0.004 (0.015)
Log of Lagged Outmigration	1.043*** (0.054)	1.064*** (0.048)	1.017*** (0.053)	1.029*** (0.053)
Log of Adjusted Savings	-0.007 (0.041)	-0.016 (0.046)	0.001 (0.033)	-0.003 (0.036)
Log of Polity Index	0.172 (0.191)	0.181 (0.197)	0.153 (0.181)	0.158 (0.187)
Log of Infant Mortality	0.118 (0.079)	0.044 (0.088)	0.088 (0.063)	0.048 (0.757)
Adjusted R^2	0.772	0.773	0.853	0.852
N	149	149	149	149
DW Statistic	1.86	1.90	1.852	1.89

Notes: ***, **, and * denotes significance at 1%, 5%, and 10% respectively.
The terms in the brackets denotes standard errors.

Table A.12: Regression Results for the Asian Sample with Infant Mortality and Telephone Lines Combinations — (Dependent Variable: Log of Emigration Flows for the Period 1970-2010)

Variables	Common Coefficient I	Cross-section Specific Fixed Effects II	Panel GLS III	Panel GLS with Cross-section fixed effects IV
Constant	0.188** (0.093)	0.116 (0.083)	0.141 (0.09)	0.124 (0.088)
Log of Battle Deaths	0.004 (0.008)	0.011 (0.01)	0.009 (0.009)	0.015 (0.01)
Log of Cereal Production	-0.005 (0.015)	0.005 (0.016)	-0.007 (0.026)	-0.004 (0.03)
Log of Natural Disasters	-0.002 (0.016)	-0.0001 (0.16)	-0.005 (0.013)	-0.003 (0.013)
Log of Lagged Outmigration	1.018*** (0.046)	1.05 (0.042)	0.996*** (0.043)	1.019*** (0.043)
Log of Polity Index	0.22 (0.175)	0.226 (0.178)	0.238 (0.165)	0.246 (0.173)
Log of Telephone lines per capita	0.012 (0.034)	0.046 (0.042)	0.021 (0.032)	0.047 (0.04)
Log of Infant Mortality	0.143** (0.066)	0.107* (0.056)	0.114** (0.052)	0.1** (0.049)
Adjusted R^2	0.772	0.775	0.847	0.848
N	172	172	172	172
DW Statistic	1.724	1.775	1.782	1.836

Notes: ***, **, and * denotes significance at 1%, 5%, and 10% respectively.
The terms in the brackets denotes standard errors.

State Space Model Methodology

We use a state-space framework to fit a model consisting of variables with different frequencies. The objective of such an exercise is to derive the fitted values of outmigration flows for the 5-year incremental period. This model was used to validate our previous findings from the panel GLS specification.

State space model has two main benefits. First, it integrates unobserved components called state variables with observable series in a single system. The second advantage of this method is that it uses a recursive algorithm called Kalman filter to recursively update the state variables. In order to illustrate this method, suppose that a time series M_t is represented as follows:

$$M_t = \mu_t + \epsilon_t, \text{ where } \epsilon_t \sim N(0, \sigma_\epsilon^2) \quad (6)$$

$$\mu_{t+1} = \mu_t + \vartheta_t, \text{ where } \vartheta_t \sim \mathrm{N}(0, \sigma_v^2) \quad (7)$$

Let the initial value of μ_t be equal to zero. In this example, M_t is the observed series with an underlying or unobserved component μ_t. In this modeling framework, the first equation is called the signal equation, while the second equation, which follows a driftless random walk, is called a state equation.

The signal equation incorporates the state variable with the observed series accounting for the measurement error ϵ_t while the state equation represents the time evolution of the state variable with innovation term ϑ_t. The purpose of the analysis is to recover or estimate the unobserved state μ_t from the observable data $\{M_t | t = 1,2, \dots \dots . . T\}$. To recover the estimates, there are three common approaches. These are filtering, prediction and smoothing. Filtering uses the information $F_t = \{M_1, \dots \dots, M_t\}$ by removing the measurement errors from the data. On the other hand, prediction uses a one-step ahead forecast of μ_t or M_t and smoothing estimates μ_t using F_t.

We use the Kalman filter algorithm to do the smoothing estimates. The main purpose of this algorithm is to recursively update the state variables when new information becomes available. The algorithm constitutes of two parts, namely, predicting and updating. In the prediction part, a one-step ahead prediction of M_t is estimated by using information from *t =1* to *t-1*. When M_t is realized at time *t,* the prediction error or innovation can be computed as $v_{t|t-1} = M_t - M_{t|t-1}$

This innovation term now contains information about the state variable μ_t, which was not captured in $\mu_{t|t-1}$ and is incorporated in estimating $\mu_t = \mu_{t|t-1} + K_t v_{t|t-1}$. K_t is called the Kalman gain or the weight assigned to the innovation.

Generating 5-year Incremental Outmigration Flows

We use decadal outmigration flows to derive the five year incremental outmigration flows using both economic and non-economic state space variables. The innovation of this paper is that the observations between every five year period were treated as missing observations. The study then took advantage of the state space models ability to handle missing observations.

Following the discussion in the previous section, let M_t be the five-yearly outmigration flows data with missing observations and μ_t be the state equation. We then assume that $\{M_t\}_{t=l+1}^{l+h}$ was missing, where $h \geq 1$ and $1 \leq l \leq T$. For $t \in \{l+1, \dots \dots, l+h\}, \mu_t$ may be expressed as a linear combination of μ_{l+1} and $\{\vartheta_j\}_{j=l+1}^{t-1}$.

Thus, for $t \in \{l+1, \dots \dots, l+h\}$, we have

$$E(\mu_t|F_{t-1}) = E(\mu_t|F_t) = \mu_{l+1|l} \quad (8)$$

$$Var(\mu_t|F_{t-1}) = Var(\mu_t|F_t) = \Sigma_{l+1|l} + (t-l-1)\sigma_v^2 \quad (9)$$

From equations (8) and (9), the Kalman filter algorithm can be used even with missing observations by equating the Kalman gain and prediction error to zero.

Each indicator that is significant in the regression equation is created one after the other, resulting in a group of 16 different models[14]. The signal equation is composed of outmigration flows on the left hand side and GDP index and other economic and non-economic indicators on the right hand side as follows:

$$\log(M_{it}) = C(1) + \begin{pmatrix} SV_{1,t} \\ \\ SV_{16,t} \end{pmatrix} \left(Ln(GDP_{i,t}), Ln(MO_{it}), \dots \dots \dots, Ln(M_{i,t-1})\right) + \epsilon_t \quad (10)$$

$$\begin{pmatrix} SV_{1,t} \\ \\ SV_{16,t} \end{pmatrix} = \begin{pmatrix} SV_{1,t-1} + \vartheta_{1,t-1} \\ \\ SV_{16,t} + \vartheta_{16,t-1} \end{pmatrix} \quad (11)$$

The error term in equation 10 is treated as a state variable and follows an AR(1) process. After the state space models were specified, the signal series was generated using the smoothed forecast from the Eviews 8.0 software package.

One-to-One Correspondence between State Space Model and Multilevel Regression Model

Following Tsimikas and Ledolter (1997), let the univariate state-space model (SSM) and time invariant regressors be defined as follows:

$$y_t = x_t'\beta + z_t'\gamma + D_t'\alpha_t + v_t$$
$$\alpha_t = \omega_t \alpha_{t-1} + \varepsilon_t \quad (1)$$

where, $t = 1, \dots, T$. The vector β is a $(p \times 1)$ vector of fixed effects, γ is a $(g \times 1)$ vector of time-invariant random effects; $x_t{}'$ is the $(1 \times p)$ vector of variables related to the fixed effects and z_t' is the $(1 \times g)$ vector of variables related to the random effects; D_t' is a $(1 \times q)$ vector and

[14] We will not report all the state space model results for the sake of brevity. They can be obtained from the authors upon request.

α_t is the $(q \times 1)$ state vector at time t; v_t is the observation noise, ω_t is a $(q \times q)$ transition matrix, and ε_t is the disturbance term in the state transition equation.

The v_t 's are uncorrelated and distributed with mean 0 and variance σ^2, the $\varepsilon_t' s$ are uncorrelated and distributed with mean 0 and variance σ_ε^2, and disturbance and observation noise terms are assumed uncorrelated. The vector of time-invariant random effects has mean 0 and covariance matrix B_{11}. The initial state $\alpha_0 = (\alpha_{10}', \alpha_{20}')'$ is partitioned where the $(q_1 \times 1)$ vector α_{10} has a diffuse prior and the $(q_2 \times 1)$ vector has a proper prior distribution with mean 0 and $(q_2 \times q_2)$ covariance matrix B_{22}.

The initial state conditions are moved into the observation equation. Let us define $\alpha_t^{\wedge} = \alpha_t - \prod_{i=0}^{t-1} \omega_{t-i} \, \alpha_0$, we can express equation (1) as follows:

$$y_t = x_t'\beta + \left(D_t' \prod_{i=0}^{t-1} \omega_{t-i} \right) \alpha_0 + z_t'\gamma + D_t'\widehat{\alpha_t} + v_t$$

$$\widehat{\alpha_t} = \omega_t \widehat{\alpha_{t-1}} + \varepsilon_t \qquad (2)$$

The specification in equation (2) is a Generalized least squares (GLS) transformation (Harvey, 1989) with the initial state α_0 fixed at value 0.

Partitioning $b_t' = D_t' \prod_{i=0}^{t-1} \omega_{t-i} = (b_{1t}' \colon b_{2t}')$, we can rewrite equation (1) as

$$y_t = x_t^{*\prime}\beta^* + z_t^{*\prime}\gamma^* + D_t'\widehat{\alpha_t} + v_t$$

$$\widehat{\alpha_t} = \omega_t \widehat{\alpha_{t-1}} + \varepsilon_t \qquad (3)$$

where, $x_t^{*\prime} = (x_t' \colon b_{1t}')$ is the $1 \times (p + q_1)$ vector at time *t* for the $(p + q_1) \times 1$ vector of "fixed" effects $\beta^* = (\beta', \alpha_{10}')'$ and $z_t^{*\prime} = (z_t' \colon b_{2t}')$ is the $1 \times (g + q_2)$ vector at time *t* for the $(g + q_2) \times 1$ vector of time-invariant random effects $\gamma^* = (\gamma', \alpha_{20}')'$. In vector form, equation (3) can be expressed as follows:

$$y = X\beta^* + Z\gamma^* + D\hat{\alpha} + v \quad (4)$$

The state-space model is thus expressed as a linear model. The "fixed" effect vector β^* is assigned a flat prior distribution. The vector of the time-invariant random effects γ^* has a N(0, B) prior distribution; the covariance matrix B consists of B_{11} and B_{22} on its diagonal, and covariance component B_{12}.

State-Space Model Validation

This study used decadal level outmigration flows and treated its five-year incremental values as missing observations. Using a state-space model, a time varying parameter model with the signal equation error term specified as an AR(1) process, the five-year predicted values of outmigration flows were obtained. The Appendix in the paper describes the formulation of the state-space model. Below is additional supporting information.

State Space Model Specification with Errors Represented as State Variable with AR(1) Process

Algorithms of the state-space model with errors represented as state variables with AR(1) process.

Sspace01: GDP and Lagged Outmigration flows

```
@signal ltmt = c(1)+ sv1*ltgdp +sv3*ltmtminusone + sv6

@state sv1 = sv1(-1) + [var = exp(c(7))]
@state sv3 = sv3(-1) + [var = exp(c(7))]
@state sv6 = c(8)*sv6(-1) + [var = exp(c(7))]
```

Sspace02: GDP, Lagged Outmigration Flows and Telephones per capita

```
@signal ltmt =  sv1*ltgdp + sv3*lttt + sv4*ltmtminusone+ sv6

@state sv1 = sv1(-1) + [var = exp(c(7))]
@state sv3 = sv3(-1) + [var = exp(c(7))]
@state sv4= sv4(-1)  + [var = exp(c(7))]
@state sv6 = c(8)*sv6(-1) + [var = exp(c(7))]
```

Sspace03: GDP, Infant Mortality and Battle Deaths

```
@signal ltmt =  sv1*ltgdp + sv3*ltmot + sv4*ltbd+ sv6

@state sv1 = sv1(-1) + [var = exp(c(7))]
@state sv3 = sv3(-1)
@state sv4 = sv4(-1)
@state sv6 = c(8)*sv6(-1) + [var = exp(c(7))]
```

Sspace04: GDP, Cereal Production, Lagged Outmigration Flows and Telephones per capita

```
@signal ltmt =  sv1*ltgdp + sv6*ltct + sv3*ltmtminusone +sv4*lttt + sv7

@state sv1 = sv1(-1) + [var = exp(c(1))]
@state sv3 = sv3(-1) + [var = exp(c(1))]
@state sv6 = sv6(-1) + [var = exp(c(1))]
@state sv4 = sv4(-1)
@state sv7 = c(8)*sv7(-1) + [var = exp(c(1))]
```

Sspace05: GDP, Infant Mortality, Infant Mortality rate, Battle Deaths, and Adjusted Savings

```
@signal ltmt = c(1)+ sv1*ltgdp + sv3*ltmot + sv4*ltbd+ sv5*ltst + sv6

@state sv1 = sv1(-1) + [var = exp(c(7))]
@state sv3 = sv3(-1)
@state sv4 = sv4(-1)
```

```
@state sv5 = sv5(-1) + [var = exp(c(7))]
@state sv6 = c(8)*sv6(-1) + [var = exp(c(7))]
```

Sspace06: Cereal production, Lagged Outmigration Flows and Telephones per capita

```
@signal ltmt = c(1)+ sv6*ltct + sv3*ltmtminusone +sv4*lttt + sv7

@state sv3 = sv3(-1) + [var = exp(c(1))]
@state sv6 = sv6(-1)
@state sv4 = sv4(-1)
@state sv7 = c(8)*sv7(-1) + [var = exp(c(1))]
```

Sspace07: Cereal production, Infant mortality rate and Lagged Outmigration flows

```
@signal ltmt = c(1) + sv1*ltct + sv2*ltmot +sv4*ltmtminusone+ sv3

@state sv1 = sv1(-1)
@state sv2 = sv2(-1)
@state sv4 =sv4(-1) + [var = exp(c(1))]
@state sv3 = c(2)*sv3(-1) + [var = exp(c(1))]
```

State Space Model Specification with Errors Represented as State Variable with AR(1) Process

Actual and smoothed forecasted values of outmigration flows for countries using GDP, Cereal Production, Lagged outmigration flows, and Telephone Lines as state space variables for the Year 1980.

Actual and Smoothed Forecasted Values of Outmigration flows for Countries using GDP, Cereal Production, Lagged Outmigration Flows, and Telephones Per Capita as State Space Variables for the Year 1980

Year	Country	Actual Outmigration flows	Predicted Outmigration flows	Difference
1980	AFGHANISTAN	1.4384	1.4384	-1.776E-15
1980	Albania	0.3511	0.3511	-1.276E-15
1980	Algeria	0.7599	0.7599	1.332E-15
1980	Angola	0.8733	0.8733	0
1980	Argentina	0.9318	0.9318	1.554E-15
1980	Armenia	0.5571	0.5571	0
1980	Australia	0.6953	0.6953	0
1980	Austria	0.6393	0.6393	9.99E-16
1980	Azerbaijan	0.8286	0.8286	8.881E-16
1980	Bahrain	1.044	1.044	4.88E-15
1980	Bangladesh	4.131	4.131	0
1980	Belarus	0.7131	0.7131	0
1980	Belgium	0.3375	0.3375	0
1980	Benin	1.292	1.292	3.552E-15
1980	Bhutan	3.149	3.149	5.32E-15
1980	Bolivia	0.6472	0.6472	0
1980	Bosnia and Herzegovina	0.7852	0.7852	0
1980	Brazil	0.863	0.863	0
1980	Bulgaria	0.556	0.556	0
1980	Burkina Faso	0.880	0.880	0
1980	Burundi	0.733	0.733	0
1980	Cambodia	1.702	1.702	0
1980	Cameroon	0.891	0.891	0

Year	Country	Actual Outmigration flows	Predicted Outmigration flows	Difference
1980	Canada	0.603	0.603	9.99E-16
1980	Cape Verde	0.768	0.768	0
1980	Central African Republic	0.732	0.732	1.11E-15
1980	Chad	0.800	0.800	1.66E-15
1980	Chile	0.993	0.993	0
1980	China	0.473	0.473	4.44E-16
1980	Colombia	1.253	1.253	1.776E-15
1980	Comoros	0.222	0.222	0
1980	Congo, Dem. Rep.	0.534	0.534	8.881E-16
1980	Costa Rica	0.775	0.775	0
1980	Cote d'Ivoire	0.6203	0.6203	8.881E-16
1980	Croatia	1.017	1.017	6.43E-15
1980	Cuba	1.686	1.686	4.88E-15
1980	Cyprus	0.863	0.863	0
1980	Czech Republic	0.602	0.602	0
1980	Denmark	0.7134	0.7134	0
1980	Djibouti	0.658	0.658	0
1980	Dominican Republic	1.338	1.338	3.10E-15
1980	Ecuador	1.011	1.011	4.662E-15
1980	Egypt, Arab Rep.	1.603	1.603	5.10E-15
1980	El Salvador	0.966	0.966	0
1980	Equatorial Guinea	2.119	2.119	0
1980	Eritrea	0.557	0.557	0
1980	Estonia	0.845	0.845	0
1980	Ethiopia	1.387	1.387	3.77E-15
1980	Fiji	1.181	1.181	2.22E-15
1980	Finland	0.768	0.768	9.99E-16

Year	Country	Actual Outmigration flows	Predicted Outmigration flows	Difference
1980	France	0.717	0.717	0
1980	Gabon	1.185	1.185	9.54E-15
1980	Gambia, The	0.704	0.704	8.88E-16
1980	Georgia	2.093	2.093	0
1980	Germany	0.679	0.679	0
1980	Ghana	1.385	1.385	0
1980	Greece	0.728	0.728	0
1980	Guatemala	1.051	1.051	4.88E-15
1980	Guinea	0.865	0.865	0
1980	Guinea-Bissau	0.618	0.618	0
1980	Guyana	1.297	1.297	8.65E-15
1980	Haiti	0.921	0.921	0
1980	Honduras	0.705	0.705	0
1980	Hungary	0.539	0.539	0
1980	Iceland	0.671	0.671	0
1980	India	0.446	0.446	8.32E-16
1980	Indonesia	0.704	0.704	0
1980	Iran, Islamic Rep.	1.033	1.033	0
1980	Iraq	0.819	0.819	0
1980	Ireland	0.876	0.876	0
1980	Israel	0.737	0.737	0
1980	Italy	0.65	0.65	0
1980	Jamaica	0.96	0.96	0
1980	Japan	0.898	0.898	0
1980	Jordan	1.085	1.085	0
1980	Kazakhstan	0.78	0.78	0
1980	Kenya	0.443	0.443	9.43E-16

Year	Country	Actual Outmigration flows	Predicted Outmigration flows	Difference
1980	Korea, Dem. Rep.	1.217	1.217	7.77E-15
1980	Korea, Rep.	0.657	0.657	8.88E-16
1980	Kuwait	0.535	0.535	0
1980	Kyrgyz Republic	2.272	2.272	0
1980	Lao PDR	1.61	1.61	0
1980	Latvia	0.876	0.876	0
1980	Lebanon	1.231	1.231	0
1980	Lesotho	0.545	0.545	0
1980	Liberia	1.521	1.521	7.1E-15
1980	Libya	0.49	0.49	0
1980	Lithuania	0.608	0.608	0
1980	Luxembourg	0.669	0.669	0
1980	Macedonia, FYR	1.487	1.487	3.99E-15
1980	Madagascar	1.783	1.783	3.55E-15
1980	Malaysia	0.78	0.78	0
1980	Mali	1.071	1.071	0
1980	Mauritania	0.661	0.661	8.88E-16
1980	Mauritius	1.654	1.654	6.21E-15
1980	Mexico	1.237	1.237	8.88E-15
1980	Moldova	0.619	0.619	0
1980	Mongolia	1.408	1.408	0
1980	Morocco	0.803	0.803	0
1980	Mozambique	0.396	0.396	-4.99E-16
1980	Myanmar	0.806	0.806	0

Year	Country	Actual Outmigration flows	Predicted Outmigration flows	Difference
1980	Namibia	1.098	1.098	2.22E-15
1980	Nepal	0.607	0.607	0
1980	Netherlands	0.672	0.672	0
1980	New Zealand	1.19	1.19	4.88E-15
1980	Nicaragua	0.855	0.855	0
1980	Niger	0.92	0.92	0
1980	Nigeria	0.534	0.534	0
1980	Norway	0.447	0.447	0
1980	Oman	0.708	0.708	0
1980	Pakistan	0.2372	0.2372	4.44E-16
1980	Panama	1.274	1.274	6.661E-15
1980	Papua New Guinea	1.424	1.424	4.44E-15
1980	Paraguay	0.70	0.70	0
1980	Peru	0.908	0.908	0
1980	Philippines	1.571	1.571	7.99E-15
1980	Poland	0.547	0.547	0
1980	Portugal	1.043	1.043	4.21E-15
1980	Qatar	0.543	0.543	0
1980	Romania	0.646	0.646	0
1980	Russian Federation	0.806	0.806	0
1980	Rwanda	0.429	0.429	9.43E-16
1980	Saudi Arabia	0.722	0.722	0
1980	Senegal	0.966	0.966	0
1980	Sierra Leone	0.95	0.95	8.88E-16
1980	Slovak Republic	0.486	0.486	8.88E-16
1980	Slovenia	0.71	0.71	0

Year	Country	Actual Outmigration flows	Predicted Outmigration flows	Difference
1980	Solomon Islands	0.624	0.624	0
1980	Somalia	0.514	0.514	0
1980	South Africa	0.662	0.662	0
1980	South Sudan	0.641	0.641	0
1980	Spain	0.643	0.643	0
1980	Sri Lanka	1.635	1.635	6.21E-15
1980	Sudan	0.741	0.741	0
1980	Suriname	1.306	1.306	4.21E-15
1980	Swaziland	0.51	0.51	9.92E-16
1980	Sweden	0.479	0.479	0
1980	Switzerland	0.749	0.749	8.88E-16
1980	Syrian Arab Republic	0.813	0.813	0
1980	Tajikistan	1.192	1.192	8.88E-15
1980	Tanzania	0.498	0.498	4.44E-16
1980	Thailand	0.39	0.39	7.21E-16
1980	Timor-Leste	1.462	1.462	0
1980	Togo	0.448	0.448	4.99E-16
1980	Trinidad and Tobago	0.901	0.901	0
1980	Tunisia	0.916	0.916	0
1980	Turkey	1.879	1.879	3.33E-15
1980	Turkmenistan	0.336	0.336	5.55E-16
1980	Uganda	0.79	0.79	0
1980	Ukraine	0.644	0.644	0
1980	United Arab Emirates	0.261	0.261	1.11E-15
1980	United Kingdom	0.758	0.758	0
1980	United States	0.872	0.872	0
1980	Uruguay	1.128	1.128	6.43E-15

Year	Country	Actual Outmigration flows	Predicted Outmigration flows	Difference
1980	Uzbekistan	1.174	1.174	3.77E-15
1980	Venezuela, RB	0.941	0.941	9.99E-16
1980	Vietnam	0.871	0.871	0
1980	Yemen, Rep.	1.147	1.147	7.54E-15
1980	Zambia	0.546	0.546	0
1980	Zimbabwe	0.564	0.564	0

Source: Authors' own computations.

APPENDIX B: ANALYSIS BILATERAL MIGRATION FLOWS

Chapter 2 presents a migration model that estimates the potential for global out migration at the country level. To complement this work, similar efforts were conducted to explore whether a model of bilateral in-migration might be developed. That is, a model that would estimate where the emigrants (from the model in Chapter 2) would migrate. Toward this goal a framework to determine the variables and tests affecting bilateral migration was constructed and implemented. The corresponding data set consisted of initially 163,350 observations across 165 countries for 6 time periods (decadal time intervals). Second, in an attempt to improve the statistical significance of the independent variables and overall explanation of changes in bilateral migration (the dependent variable) another analysis was developed for 160 countries. This was done to eliminate several countries that had little to no data available for the independent variables.

Both analyses did not adequately capture the bilateral migration behaviors that were expected returning an adjusted R^2 of 0.638 for the 165 country analysis, and 0.367 for the 160 country analysis. These results indicate that only 63.8% and 36.7% of the variation in bilateral migration can be explained from the regression equation analyses for the 165 and 160 country analyses, respectively.

Determinants of Bilateral Migration

This study focuses on the determinants of bilateral migration between country pairs for a sample of 165 countries. We have a total 165*165*6 = 163350 observations in our sample for both the independent and dependent variables, namely bilateral migration flows from country *i* to country *j* (ΔlnM_{ijt}); GDP differences between each country pair $(\Delta lnGDP_{ijt})$, polity differences ($\Delta\ lnS_{ijt}$), cereal production differences (ΔlnC_{ijt}), adjusted savings differences $(\Delta ln\ PTNEW_{ijt})$, telephone infrastructure difference $(\Delta lnTT_{ijt})$, natural disaster (Δlnd_{ijt}) and battle deaths differences $(\Delta lnBD_{ijt})$, and infant mortality differences $(\Delta lnIM_{ijt})$. All the variables are normalized between 0 and 1 and transformed into logarithmic differences.

Model Specification

Our model is of the following form:

$$\Delta lnM_{ijt} = f(\Delta lnGDP_{ijt}, \Delta\ lnS_{ijt}, \Delta lnC_{ijt}, \Delta ln\ PTNEW_{ijt}, \Delta lnTT_{ijt}, \Delta lnd_{ijt}, \Delta lnBD_{ijt}, \Delta lnIM_{ijt}) \quad \text{(B.1)}$$

where the differences are between country pairs *i* and *j*. Both the notations of dependent and independent variables remain the same.

We expect that the lagged bilateral migration will have a positive effect on current migration flows, GDP differences will have a negative effect on bilateral migration flows, differences in savings will have a negative effect on bilateral migration flows, greater number of disasters between country pairs will have a positive effect on bilateral migration flows, and the greater number of battle deaths between country pairs will have a positive effect on bilateral migration flows. Infrastructure differences, polity differences and cereal production are likely to have ambiguous effect on bilateral migration flows.

Dataset

The data originated in the Excel spreadsheet “Yearly factor dat.xlsx,” which is a compilation of World Bank generated information. Annual data for each of the 165 countries were averaged into five-year averages. Data streams of Cereal, Infant Mortality, GDP, Savings, Phone, Polity, Battle Deaths, and Disaster then became the input for the “difference” processing.
Difference processing begins with a “county of interest”. The difference between two countries is expressed as a proportion of the country of interest as per this formula, (country value – country of interest value) / (country of interest value). These difference proportions may be positive, negative, zero, minus one (-1), or undefined (Div0 error). Positive, negative, and zero values are acceptable but negative one and the Div0 error actually represent a zero value. These cases were filtered out and converted to zero. This produces a 165x165 = 27,225 element matrix for each of the 6, decadal, 5 year averages. As the human migration data was only available as “migrant percent” at decade intervals we processed the data streams corresponding to those times, thus for example the difference became (data value at 1970-data value at 1960). This produces 5 lagged matrices of 165x165 elements, (1970-1960), (1980-1970), (1990-1980), (2000-1990), and (2010-2000). The data for each country within these five matrices was then normalized between 0 and 1. So the column representing Afghanistan and its other 164 values was normalized between 0 and 1. Then the column representing Albania was normalized between 0 and 1, etc. Finally the 5 lagged matrices for the specific data stream (Cereal, Mortality, etc.) were combined into a vector of 165x165x5 = 136,125 elements.

Dependent Variable: Migration Rate

- Data Source-World Bank Bilateral Estimates of Migrant Stocks. Data available at decade frequency from 1960-2010.
- Data available for 165 countries
- Took difference in migration between country pairs *i* and *j*. This variable is of dimension 165*165.
- Scaled the variable between 0 and 1

Independent Variable: Disaster Impact

- Measure of environmental security, e.g., do I have the basic services that I need
- Data Source-World Bank World Development Indicators. Data available at annual frequency from 1960-2012. Availability of data varies by variable
- Data available for 165 countries
- The disaster differences between country pair *i* and *j* are chosen and this variable is of dimension 165*165.
- Scaled this variable between 0 and 1

- Variable should be positively correlated. As disasters increase so should bilateral migration flows

Independent Variable: Cereal Production

- Measure of food security, e.g., can I feed myself
- Data Source-World Bank World Development Indicators. Data available at annual frequency from 1960-2012. Availability of data varies by variable
- Data available for 165 countries
- Took the per-capita cereal production differences between country pairs *i* and *j*. This variable is of dimension 165*165.
- Scaled this variable between 0 and 1
- Variable should be negatively correlated. As cereal production increases bilateral migration flows should decrease

Independent Variable: GDP

- Measure of financial security, e.g., do I make enough to meet my basic needs
- Data Source-World Bank World Development Indicators. Data available at annual frequency from 1960-2012. Availability of data varies by variable
- Data available for 166 countries
- Took the difference of GDP between country pairs *i* and *j*. This variable is of dimension 165*165.
- The variable is scaled between 0 and 1
- Variable should be negatively correlated. As GDP increases bilateral migration flows should decrease

Independent Variable: Mortality

- Measure of health security, e.g., do I have access to good and affordable health care
- Data Source-World Bank World Development Indicators. Data available at annual frequency from 1960-2012. Availability of data varies by variable
- Data available for 165 countries
- Took average mortality rate (children under 5) over each decade, for each country and normalized it by the average population over the 10 year period
- Scaled the infant mortality variable between 0 and 1
- Variable should be positively correlated. As mortality increases so should bilateral migration

Independent Variable: Battle Deaths

- Measure of personal security, e.g., is my life in danger from crime or war
- Data Source-World Bank World Development Indicators. Data available at annual frequency from 1960-2012. Availability of data varies by variable
- Data available for 165 countries

- Summed total battle deaths over each decade, for each country and normalized it by the average population over the 10 year period
- Took the difference of this variable between country pairs *i* and *j*. This variable is of dimension 165*165.
- The variable was normalized between 0 and 1.

Independent Variable: Polity
- Measure of social security, e.g., can I depend on my government
- Data Source-need to check. Data available at annual frequency from 1960-2012. Availability of data varies by variable
- Data available for 165 countries
- Took the difference of the polity score over each decade, for each country
- Scaled the variable between 0 and 1
- Variable should be negatively correlated. As polity increases out-migration should decrease

Table B.1: Descriptive Statistics

	LN_DISASTERS	LN_MORTALITY	LN_PHONE	LN_POLITY	LNBD	LNCEREAL	LNGDP
Mean	0.039713	0.195800	0.123695	0.101875	0.004025	0.038324	0.068948
Median	0.000106	0.134784	0.033470	0.000000	5.00E-12	0.008464	0.010558
Maximum	0.693147	0.693147	0.693147	0.693147	0.693147	0.693147	0.693147
Minimum	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
Std. Dev.	0.104330	0.186536	0.176086	0.233463	0.036636	0.089736	0.134056
Skewness	3.697234	0.677282	1.582895	1.883931	14.14139	4.587571	2.629784
Kurtosis	18.05128	2.238805	4.465144	4.598079	229.7552	27.75049	9.594370
Jarque-Bera	1914047.	16432.05	82824.36	114009.0	3.55E+08	4742389.	484256.3
Probability	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
Sum	6487.168	31983.95	20205.59	16641.35	657.4202	6260.234	11262.73
Sum Sq. Dev.	1778.007	5683.816	5064.828	8903.333	219.2427	1315.366	2935.540
Observations	163350	163350	163350	163350	163350	163350	163350

Table B.2: Correlation Matrix

	LN_DISASTERS	LN_MORTALITY	LN_PHONE	LN_POLITY	LNBD	LNCEREAL	LNGDP
LN_DISASTERS	1.000000	0.212556	-0.142381	-0.018243	0.104601	0.022331	-0.146449
LN_MORTALITY	0.212556	1.000000	-0.535829	0.105843	0.057412	-0.101389	-0.390477
LN_PHONE	-0.142381	-0.535829	1.000000	-0.069193	-0.032894	0.112146	0.721097
LN_POLITY	-0.018243	0.105843	-0.069193	1.000000	0.041943	-0.022387	-0.089847
LNBD	0.104601	0.057412	-0.032894	0.041943	1.000000	-0.011882	-0.041510
LNCEREAL	0.022331	-0.101389	0.112146	-0.022387	-0.011882	1.000000	0.144211
LNGDP	-0.146449	-0.390477	0.721097	-0.089847	-0.041510	0.144211	1.000000
LNMIGRATION	-0.025402	-0.038007	0.114070	0.004612	-0.003720	-0.037701	0.104322

Methodology: Time Series Analysis

The goal of this paper is to estimate bilateral migration flows between pairs of countries using our explanatory variables. However, in order to prevent estimating a spurious regression, the time series properties of the variables of study are determined before the estimation procedure is chosen. Both the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) unit root tests are performed on each series in order to determine their order of integration. Results from the unit root tests would determine the procedure to be used to estimate the bilateral migration flows. For instance, if all series are integrated of order 0, then OLS may be used. In contrast, if the series are unit root non-stationary, then OLS would generate a spurious regression and an alternative method to OLS should be explored.

The ADF unit root test involves estimating equation (B.2), and then, testing the null hypothesis of a unit root, H_0: $\alpha = 0$ versus the alternative of a stationary process, H_1: $\alpha < 0$. The test is based on the typical t-ratio for α (Dickey and Fuller, 1979). However, the t-statistic does not follow the t-distribution under the null and thus the critical values are simulated for each regression specification and the sample size (Mackinnon, 1996).

$$\Delta y_t = \alpha y_{t-1} + x_t'\delta + \sum_{i=1}^{p} \Delta y_{t-i} + \varepsilon_t \qquad \text{(B.2)}$$

x_t consists of exogenous regressors that can include a constant term only, a constant and a trend, or none. Δy_{t-i} include terms that correct for higher-order correlation.

The PP unit root test involves estimating a non-augmented version of regression (B.2) i.e. without the lagged difference terms. The PP unit root test uses a non-parametric method to control for serial correlation under the null hypothesis. H_0 and H_1 are the same as in the ADF test; however the PP unit root test is based on its own statistic and the corresponding distribution (Phillips and Perron, 1988).

Based on the unit root tests in table B.3, we find that all the variables whether in levels or in first differences and with a drift and drift and trend term satisfy stationarity. Thus, we can estimate our model using dynamic OLS. The Stock-Watson OLS can be specified as follows:

$$Y_t = \beta_0 + \vec{\beta}\, X + \sum_{j=-q}^{p} \vec{d_j}\, \Delta X_{t-j} + u_t \quad (3)$$

where Y_t is the dependent variable bilateral migration flows between country *i* and *j*
X denotes the matrix of explanatory variables
$\vec{\beta}$ is the cointegrating vector i.e. it represent the long-run cumulative multipliers
p is the lag length
q is the lead length

Lag and lead terms are included in the DOLS regression to make the stochastic error term independent of all past innovations in stochastic regressors. Finally, we also perform the residuals of the estimated DOLS regression in order to test whether it is a spurious regression. These tests are performed using Eviews 8.

Table B.3: Unit Root Tests: ADF and PP Tests

Test Statistic by Unit Root Test	Variables		Ln M (lnmigration)	Ln GDP (lngd)	Ln C (lncereal)	Ln Mot	Ln TT	Ln S (lnsavings)	Ln D (ln_disasters)	Ln PTNEW (ln_polity)	Ln BD (lnbd)
	ADF: drift	Level	**-286.03*****	**-254.87*****	**-291.14*****	-337.87***	-242.31***	**-106.71*****	**-259.34**	**-37.37*****	**-255.11**
		1st difference	**-492.23*****	**-437.17*****	-486.9***	-491.53***	-440.76***	**-475.2*****	**-475.33*****	**-495.83*****	**-481.56*****
	ADF: drift and trend	Level	**-286.03*****	-342.66*** -263.2245 ***	-294.61***	-338.43***	-265.67***	**-109.47*****	**-267.89**	**-48.65*****	**-255.13*****
		1st difference	**-492.23*****	**-437.17*****	-486.89***	-491.53***	-440.76***	**-475.2*****	**-475.33*****	**-495.83*****	**-481.56*****
	PP: drift	Level	-401.83***	**-691.05*****	-524.27***	-626.07***	-883.27***	-106.71***	-259.34***	-37.37***	-609.54***
		1st difference	-24944.06***	**-4515.17*****	-7393.12***	-3657.12***	-4618.14***	-2057.58***	-8201.26***	-4568.91***	-14064.55***
	PP: drift and trend	Level	-401.83***	**-545.54*****	-461.67***	-616.38***	-586.48***	-623.68***	-624.80***	-280.78***	-609.13***
		1st difference	-24943.92***	**-4515.56*****	-7393.09***	-626.07***	-4618.12***	-2057.57***	-8201.22***	-4568.99***	-14064.44***

Notes: H_0 : Unit root process for ADF and PP. *, **, *** refers to the rejection of H_0 at 0.1, 0.05 and 0.01 significance levels respectively. Number of lags in ADF tests was based on the modified AIC and the optimal lag was 1 for each series.

Dynamic OLS Estimation Results

Results from unit root tests show that all series under study are unit root non-stationary (See Table B.3). In particular, all specifications of ADF and PP tests cannot reject the null hypothesis of a unit root process. Given that all the series in the bilateral migration flows are unit root non-stationary, then the cointegrating regression to be estimated is as follows:

$$Ln\, M_{ijt} = \beta_0 + \beta_1\, lnGDP_{ijt} + \beta_2\, lncereal_{ijt} + \beta_3 lnM_{ijt-1} + \beta_4\, trend + \beta_5\, lnMOT_{ijt-1} + \beta_6\, lnS_{ijt-1} + \beta_7\, lnTT_{ijt-1} + \sum_{j=-q}^{p} \overrightarrow{d_1 \Delta}\, lnGDP_{t-j} + \sum_{j=-q}^{p} \overrightarrow{d_2 \Delta}\, lncereal_{t-j} + U_t$$
(B.4)

Given that all the series in the Bilateral migration flows are unit root non-stationary, then the cointegrating regression equation is estimated as given by equation (B.4). This specification is the one commonly used in the literature on bilateral migration flows. Table B.4 presents the DOLS estimation results. The number of leads and lags were selected according to the Hannan-Quinn Criterion.

The estimation results show that for the full specification GDP differences between country *i* and *j,* cereal production differences between country *i* and *j,* lagged bilateral migration flows are all significant determinants of bilateral migration flows. However, the estimated coefficients are all attenuated relative to our outmigration flows model. In other words, a 1 percent difference between GDP between country *i* and country *j,* will on an average, lead to 0.0006 percent difference in bilateral migration flows. Similarly, a 1 percent difference in cereal production between country *i* and country *j,* will on an average, lead to 0.019 percent difference in bilateral migration flows between country pairs. In addition, a 1 percent difference in lagged bilateral migration flows will lead to a 0.61 percent difference in migration flows between country pairs. From Table B.3, all the specifications of ADF and PP unit root tests conclude that the residuals are stationary. Thus, our regression model in DOLS is not a spurious regression.

Table B.4: DOLS Estimation Results for Bilateral Migration Flows

Estimated Coefficients		t-values
β_0	0.0006*** (0.0002)	3.012
β_1 ($lnGDP_{ijt}$)	-0.003** (0.001)	-1.936
β_2 ($lncereal_{ijt)}$	0.019*** (0.003)	5.048
β_3($lnmigration_{ijt-1}$)	0.612*** (0.013)	44.06
β_4 (lntrend)	1.32E-09 (1.8E-09)	0.735
β_5 ($lnmortaility_{ijt-1}$)	-0.0001 (0.0002)	-0.629
β_6 ($lnsavings_{ijt-1}$)	-0.0002* (0.0001)	-1.664
β_7 ($lnphone_{ijt-1}$)	0.0006*** (0.0001)	3.476
Adj R^2	0.638	---
SE	0.02	---
DW Statistic	2.09	---
Leads	19	---
Lags	68	---

Notes: ***, **, * denotes statistical significance at 0.1, 0.05, and 0.01 significance levels respectively. t-statistics appears in column 3; leads and lags are selected according to the Hannan-Quinn Criterion.

Table B.5: Unit Root / Stationarity Tests on Residuals from DOLS Estimation Results

		ADF drift	ADF Drift and Trend	PP drift	PP Drift and Trend
DOLS residuals	Level	**-269.99*** **	**-269.99*** **	**-390.61*** **	**-390.60*** **
	1st Difference	**-461.17*** **	**-461.16*** **	**-19570.18*** **	**-19570.45*** **

Notes: H_0 : Unit root process for ADF and PP. *, **, *** refers to the rejection of H_0 at 0.1, 0.05 and 0.01 significance levels respectively.

Conclusions – 165 Country Analysis

The results of the determinants of bilateral migration flows strongly suggest that cereal production differences, GDP differences between countries, and lagged bilateral migration flows are significant, although the estimated elasticities are much lower than our model of outmigration flows. This may be because of the additional temporal dimension of differences between pairs of countries between the dependent and independent variables in the model. Our results also suggest that infrastructure differences between pairs of countries are positive and significant in explaining bilateral migration flows. This result may suggest that amenity differences between two areas is an important push or pull factor for migration to occur. Further research in this area should include amenity driven factors, such as unemployment rates between countries. Unfortunately, such information is missing for many regions.

160 Country Analysis

A key challenge in the bicountry migration analysis is obtaining sufficient data. To eliminate countries that had several zero values in cereal, mortaility, CGP and Phone lines which suggest non-reported data all the country's data was analyzed. Any countries with more than 11 non-reported data entries were eliminated to ideally improve the quality of the dataset where possible. Figure B.1 illustrates all 165 countries where the frequency of each country's zero values were summed to help identify countries with

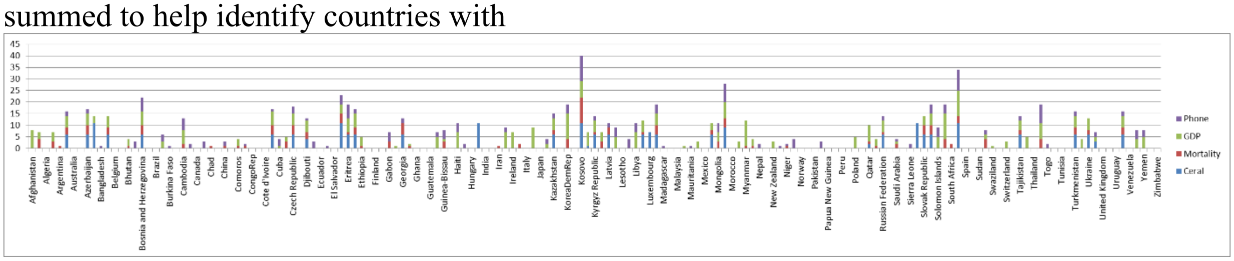
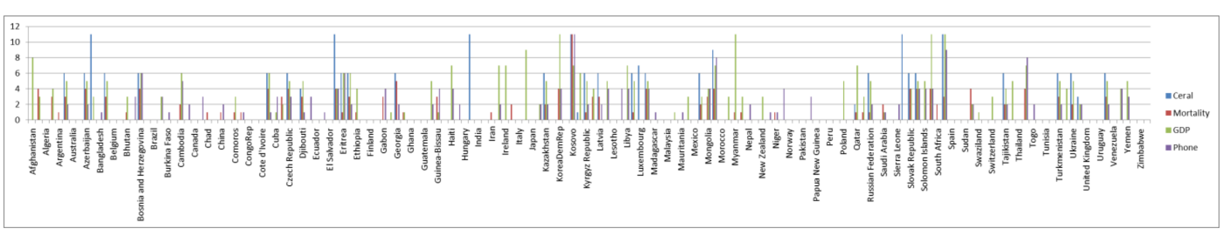


Figure B.1: Number of zero data entries for each of the 165 countries in the bilateral migration analysis.



Removing the counties with more than 11 zero data entries offers a more focused dataset to run the similar analysis on as that run on the dataset with 165 countries.

Table B.6: Unit Root Tests: ADF and PP Tests

Variables											
Test Statistic by Unit Root			LNMIGRATION	LNGDP	LNSAVINGS	LNCEREAL	LN_POLITY	LN_PHONE	LN_DISASTERS	LNBD	LN_MORTALITY
	ADF: drift	Level	-277.7009***	-252.8610***	-115.0666***	-285.2219***	-273.8779***	-240.2951***	-249.6888***	-251.4914***	-254.4844***
		1st difference	-476.8312***	-425.6912***	-448.9702***	-471.6877	-470.9051***	-422.5687***	-465.5525***	-522.7692***	-469.2733***
	ADF: drift and trend	Level	-277.7004***	-264.1034***	-118.3877***	-288.5507***	-276.3156***	-265.7953***	-258.4508***	-251.5569***	-255.5182***
		1st difference	-476.8297***	-425.6898***	-448.9688***	-471.6862***	-470.9035***	-422.5673***	-465.5509***	-522.7675***	-469.2718***
	PP: drift	Level	-385.1997***								
		1st difference	-697.3089***								
	PP: drift and trend	Level	-385.1988***								
		1st difference	-697.3065***								

Notes: H_0 : Unit root process for ADF. *, **, *** refers to the rejection of H_0 at 0.1, 0.05 and 0.01 significance levels respectively. Number of lags in ADF tests was based on the modified AIC and the optimal lag was 1 for each series.

Dynamic OLS Estimation Results

Results from unit root tests show that all series under study are unit root non-stationary (See Table B.6). In particular, all specifications of ADF tests cannot reject the null hypothesis of a unit root process. Given that all the series in the bilateral migration flows are unit root non-stationary, then the cointegrating regression to be estimated is as shown in Equation B.4. Table B.7 presents the DOLS estimation results. The number of leads and lags were selected according to the Hannan-Quinn Criterion.

From Table B.6, all the specifications of ADF unit root tests conclude that the residuals are stationary. Thus, our regression model in DOLS is not a spurious regression. The dynamic ordinary least squares estimator was employed to determine how significant the relationship is between the independent variables and the dependent variable, migration between countries. Table B.7 illustrates the results of this estimation.

The results in table 7 illustrate a relatively low adjusted R^2 which is to say approximately 38.6% of the changes in the bilateral migration can be explained by the types of independent variables, and their formulation on the right hand side of the DOLS equation. Additionally, the independent variables based on the access to phones, savings rate, and cereal production do not have a significant relationship to bilateral migration at the 90% level of confidence.

Table B.7: DOLS Estimation Results for Bilateral Migration Flows
(Note: Results are for the "linear trend" specification option for the DOLS equation estimation).

Estimated Coefficients		t-values (Prob.)
β_0	0.004613	19.43907 (0.0000)
β_1 (lnGDP_{ijt})	0.005517	2.074235 (0.0381)
β_2 (ln$cereal_{ijt)}$	0.009866	1.610939 (0.1072)
β_3(ln$migration_{ijt-1}$)	0.665662	24.05368 (0.0000)
β_4 (lntrend)	-1.45E-08	-4.522192 (0.0000)
β_5 (ln$mortaility_{ijt-1}$)	-0.011781	-10.37155 (0.0000)
β_6 (ln$savings_{ijt-1}$)	-0.000255	-0.337749 (0.7356)
β_7 (ln$phone_{ijt-1}$)	0.002013	0.834817 (0.4038)
Adj R^2	0.385804	---
SE	0.027893	---
DW Statistic	1.988096	---
Leads	74	---
Lags	33	---

Notes: ***, **, * denotes statistical significance at 0.1, 0.05, and 0.01 significance levels respectively. t-statistics appears in column 3; leads and lags are selected according to the Hannan-Quinn Criterion.

The estimation results show that for the full specification GDP differences between country *i* and *j,* lagged bilateral migration flows, and mortality differences are all significant determinants of bilateral migration flows. Following a similar interpretation as in the 165 country-based analysis with updated results, a 1 percent difference between GDP between country *i* and country *j,* will on an average, lead to 0.0055 percent difference in bilateral migration flows across the decadal timeframe under investigation. Similarly, a 1 percent difference in lagged migration by one period (a decade) between country *i* and country *j,* will on an average, lead to 0.6657 percent difference in bilateral migration flows between country pairs. In addition, a 1 percent difference in mortaility rates will lead to a -0.0118 percent difference in migration flows between country pairs.

First Differences DOLS Analysis

Another DOLS using first differences equation was developed and analyzed given the relatively weak significance of the results using the DOLS equation specification developed for the results in Table B.7. Table 8 illustrates those results.

Table 8: DOLS First Differences Estimation Results for Bilateral Migration Flows
(Note: Results are for the "linear trend" specification option for the DOLS equation estimation).

Estimated Coefficients		t-values (Prob.)
β_0	0.005372	17.43492 (0.0000)
β_1 (lnGDP_{ijt})	-0.002011	-0.558695 (0.5764)
β_2 (ln$cereal_{ijt)}$	0.004390	0.524088 (0.6002)
β_3(ln$fdiffmigration_{ijt-1}$)	0.000275	0.304041 (0.7611)
β_4 (lntrend)	0.005372	17.43492 (0.0000)
β_5 (ln$mortaility_{ijt-1}$)	-0.000665	-0.812512 (0.4165)
β_6 (ln$savings_{ijt-1}$)	-5.84 E-05	-0.099210 (0.9210)
β_7 (ln$phone_{ijt-1}$)	0.000576	0.714485 (0.4749)
Adj R^2	0.114586	---
SE	0.033480	---
DW Statistic	1.982258	---
Leads	73	---
Lags	73	---

Notes: ***, **, * denotes statistical significance at 0.1, 0.05, and 0.01 significance levels respectively. t-statistics appears in column 3; leads and lags are selected according to the Hannan-Quinn Criterion. First differences were used for the migration, mortaility , savings and phone independent variables.

The results using the first differences in Table B.8 illustrate a lower adjusted R^2 at 0.11 than in those in Table B.7 without the first differences. Additionally, none of the independent variables are statistically significant at the 90%, or even approximately 60% level of confidence save for the trend variable. These results suggest a greater degree of equation misspecification using the first differences than those without the first differences. This may be due to several reasons including missing or misspecified variables, lack of data across the time period (decadal) and set of countries under analysis (160 countries), type of regression estimator (DOLS) and lack of data underlying the analysis (the use of non-zero data to represent missing data via an extremely low value in place of missing data). The latter may be skewing the panel data to the point that determining the relationship between the dependent variable (bilateral migration) and the independent variables (GDP, cereal production, population migration, mortaility rates, savings rates, access to telephone technology) proved to be unattainable from an acceptable statistically-significant standpoint.

References

Guttirrez, P. C., 2010. Dynamic OLS estimation of the U.S. import demand for Mexican Crude Oil, Available at: https://ideas.repec.org/p/pra/mprapa/30608.html

Ruyssen, I., Everaert, G. and G. Rayp, 2014. Determinants and dynamics of migration to OECD countries in a three-dimensional panel framework, Empir Econ., 46: 175-197.

APPENDIX C: SPILLOVER MODEL MATHEMATICAL DESIGN

Overview:
Since this work focuses on less developed or developing countries, there are limited historical data that can be used to parameterize the model and assess confidence in its results. The design of this model therefore makes a trade-off to favor a "logic-rich, data-poor" approach. For credibility, the equations and theory within the model have to be testable, and are therefore constrained to those formulations that allow comparison to whatever historically quantified data is available. Functional forms used in the model are causal to allow realistic policy testing.

To allow for the consideration of climate effects, the modeling approach assumes a 70-year time horizon running from 1990 to 2060. The model contains the following key elements/features, with all non-policy elements endogenous and dynamic.

- The model considers a population of Malians, with sub-populations distinguished by location of residence (split between five regions: urban Mali, rural Mali, neighboring countries, United States, and rest of the world), gender (male and female), labor type (skilled or common), and age category (0-14, 15-64, and 65+).
- Migration and violence behaviors are determined using a cognitive sub-model that incorporates principles from psychological theory and empirical studies.
- All behaviors reflect institutional, traditional, and cultural differences among populations.
- Economic modeling includes gross regional product (GRP), wages, income, and employment as functions of governance, technology, investment, land, resources, water, food, disease, remittances, and climate change.
- Food, water, and resource availability are based on relatively simple supply and demand models that act placeholders for more sophisticated models.
- Ecological damage effects can be added to the model, but are not included in this version.
- Climate effects include temperature, precipitation, and extreme event designations.

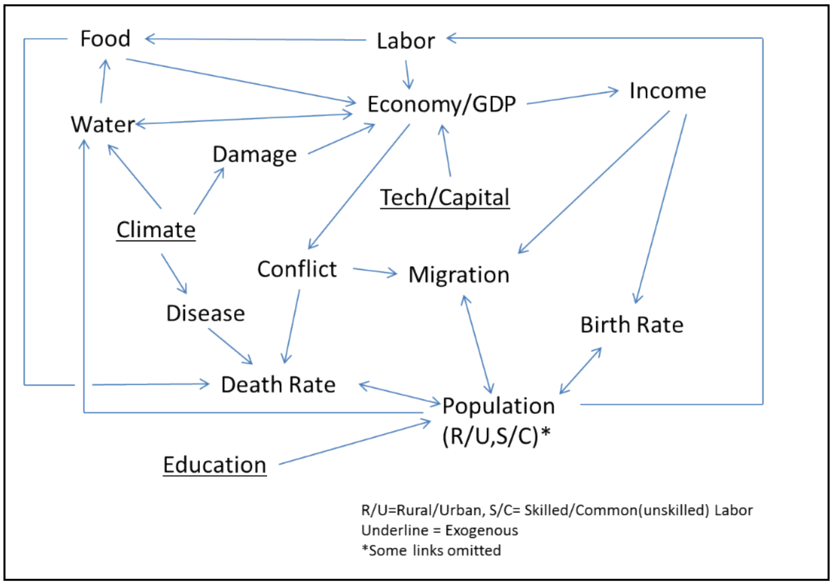


Figure C.1 Overview of the basic model interactions and considerations.

Organization

The model has population, economy, labor, resources, disease, food, water, violence, and migration sectors. Migration and violence are co-dependent. Climate is assumed to be probabilistic and sampled from the CMIP5 data set.[15] Model output is therefore an ensemble-coverage rather than a point-prediction. Intervention policies reduce the implied risk by limiting outlier (user-defined, unacceptable) impacts. All equations have array-variables and assume a differential-algebraic form. The array designations, for translating the equations are shown below.

Sets:
i = region (u,r,n,a,w)
l = labor (c,s)
g = gender (f,m)
v = age (y,m,o)

Subscripts:
x = country of interest [u,r]
u = urban region
r = rural region
n = neighboring region
a = U.S. region
w = rest-of-world region
c = common labor

[15] http://cmip-pcmdi.llnl.gov/cmip5/

s = skilled labor
f = female
m = male
y = young age (0-13)
p = productive age (14-62)
o = old age (62+)

Figure C.2 shows the flow logic for migration and the regional distinctions.

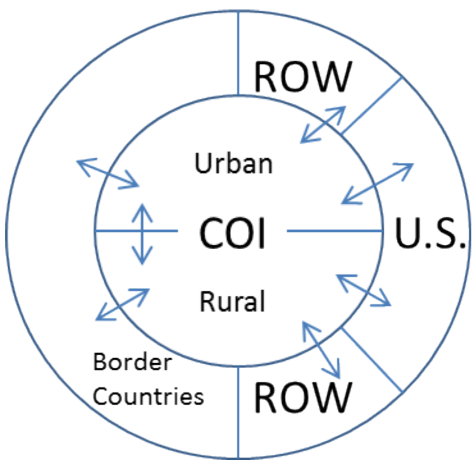


Figure C.2: Regional Distinction and Migration Flows.

Economy

The economic dynamics are based on the Cobb-Douglas (CD) construct that empirically corresponds to developing country evolution. Note that the CD formulation is inadequate for market and fiscal policy simulation (Miller 2008).

The Potential Gross Regional Product PGRP(i)

For regions [u,r,n]

Technology is adjusted for a generalized Solow production function.

$$PGRP_i = PGRP_0 * \left(\frac{LS_{i,s}}{LS_{0i,s}}\right)^{LsF_i} * \left(\frac{LS_{i,c}}{LS_{0i,c}}\right)^{LcF_i} * \left(\frac{C_i}{C_{0i}}\right)^{CF_i} * \left(\frac{TK_i}{TK_{0i}}\right)^{(LcF_i + LsF_i)}$$

$$LS_{i,l} = \sum_{j,g} POP_{i,j,l,g,p} * LPR_{j,l,g}$$

For regions [w, a]:

$$PGRP_i = exo.f(time)$$

For all regions:

$$C_{i,t} = Integral\left(INV_{i,t} - DEPR_{i,t}\right)$$

Depreciation

$$DEPR_{i,t} = C_{i,t}/PL$$

Investment

$$INV_{i,t} = CAP_{i,t-1} * IF_{i,t} * GG_{i,t}{}^{\upsilon_i} * GI_{i,t}{}^{\eta_i}$$

Technology is calibrated to set RGRP to data.

PL = Physical lifetime = 20 years
η_i = Rule of law impact on investment (assume 0.0)
υ_i = Corruption impact on investment (assume 0.0)
IF = Investment fraction
CF = Capital fraction =1.0-Lcf-LsF
C_i = Capital
$E_{i,l}$ = Employees
GG = Governance effectiveness (Corruption Index data)
GRP_i = Gross regional product (rural = agriculture, Urban = all less agriculture, regional = all)
LcF = Common labor fraction (from data)
$LPR_{j,l}$ = Labor participation rate
LsF = Skilled labor fraction (from data)
$POP_{i,j,l,g,v}$ = Population
TK_i = Technology

Y_0 = Any "0 subscript" parameter is an initial condition derived from historical data.
$\delta_{i,j}$ = complimentary Kronecker Delta function: $\delta_{i,j}$=1 for $i \neq j$; $\delta_{i,j}$=0 for i=j
λ_i= Capital Growth =(historical country growth rate)$^{1/2}$ as approximation or as scenario (Used for reference only)
μ_i = Technology growth = (historical country growth rate)$^{1/(2*CFi)}$ as approximation, or as scenario
{Exponents are general functional forms with actual values different in each equation, despite identical names}

Realized GDP (i) [again assumes a CD construct – log-linear formulation]

$$RGRP_i = PGRP_i * EC_i{}^{\alpha_i} * EG_i{}^{\beta_i} * ELs_i{}^{\gamma_i} * ELu_i{}^{\delta_i} * RA_i{}^{\epsilon_i} * GG_i{}^{\mu_i} * GI_i{}^{\sigma_i}$$

$$EC = CT^{\alpha} * WA^{\beta} * EE^{\gamma}$$
$$EG = CT^{\alpha} * WA^{\beta} * EE^{\gamma}$$
$$EL = CT^{\alpha} * WA^{\beta} * FA^{\gamma}$$

$GG_i = exo.f(time)$
$GI_i = exo.f(time)$

For EC
α=-0.01 Van Art 2004
β=-1.0 (Chu study) Backus 2013
γ= 0.0 (Lowry)
For EG
α=-0.02 McCarl 2008
β=-0.04 McCarl 2008
γ= 0.0 (Lowry)
For EL
α= -4.3 Dunne 2013 for unskilled "outside" workers.[16]
β=1.28 Kenefick 2007
γ= 0.73 Strauss 1985

LPR from EVFA.xls Totals tab.
EC = Effective capital
ELs, Elu = Effective labor (skilled, unskilled)
EG = Effective land
WA= Water availability
EE = Extreme events (From CMIP5 data)
CT = Climate temperature ratio change =current/"normal"= temperature/14.0
GI=Government infrastructure/services (Competitive Index data – including rule of law)

RA = Resource availability
FA = Food availability

ε_i = 0.0
μ_i=0.01536 (From Fayissa and Nsaih 2010: rescaled to "100" scale)
σ_i=0.04273(From Fayissa and Nsaih 2010: rescaled to "100" scale, but using Govt Effectiveness as proxy)

Labor

Labor dynamics follow from the CD approach, except that wages are those due to "institutionally allowed" supply and demand dynamics rather than the conventional, optimal, equilibrium values.

$$LV_{i,l} = \frac{RGDP_i}{RGDP_{0i}} * LV_{0i,l}$$

[16] It is the current -4.3 for unskilled labor in Africa, -2.15 for skilled labor in Africa and for skilled/unskilled in ROW. (I checked with our "experts" here who have lived in Africa, And I am assuming 50% of unskilled worker are in climate controlled environments in ROW). -1.1 for skilled labor in the U.S. and -2.15 for unskilled in the U.S. (Assumed 25% of skilled are either outside like construction, or in limited climate control, like warehouses -- but 50% for unskilled).

$$E_{i,l} = (\frac{TK_{0,i}}{LV_{0i,l}} * \frac{LV_{i,l}}{TK_i})^{1/(1+\theta)} * \left(LS_{i,l} * \left(1 - UER_{0i,l}\right)/E_{0i,l}\right)^{\theta/(1+\theta)} * E_{0i,l}$$

$$LW_{i,l} = Integral(ILW_{i,l} - LW_{i,l}))/WAT_i$$

We are using the Generalized Solow CD, and indicated wages then increase with technology and supply-demand responses.

$$ILW_{i,l} = (\frac{E_{i,l}}{LS_{i,l} * (1 - UER_{0i,l})})^{\theta} * (\frac{TK_i}{TK_{0i}}) * LW_{0i,l}$$

For $LW_{0i,j}$ see below.

θ = 0.24 from REMI US model - Treyz 1993
UER0 from EVFA.xls Totals tab.

$$UER_{i,l} = 1.0 - E_{i,l}/LS_{i,l}$$

For i = r,u

$$WI_{i,l} = (LW_{i,l} * POP_{i,i,l,m,p} + \sum_j LW_j * POP_{j,x,l,m,p} * IRF_{i,j})/POPx$$

For i = n,a,w

$$WI_i = LW_i$$

CE = Capital effectiveness
E = Employment
ILW = Indicate labor wages
IRF=International remittance fraction
LS = Labor supply (only for productive age-p)
LV = Labor value
LW = Labor wages
UER = Unemployment rate
UER_0 = Initial unemployment rate
WAT = Wage adjustment time
WI = Wage income
WR = Wage ratio

Wages are a proportional (comparative affect), use of a single wage group per skill class assumes that discrimination to women and immigrants is relatively uniform across regions. Neither data nor theory supports a greater breakdown of wages.

IRF = 0.10
WR_i= from GTAP; limiting assumption that it is constant over time

	from GTAP 6490.pdf[17] wage ratio (S/C - skilled to unskilled
Rural Mali	4.25
UrbanMali	4.25
Neighbors	4.25
USA	1.38
ROW	1.16

Resource

The resource sector assumes simple regeneration. Usage is limited by minimum extraction times/maximum extraction rates (MRPR).

$$RA_i = MIN(1, RC_i/AU_i)$$

$$AU_i = Integral((RU_i - AU_i)/RIT_i)$$

$$R_i = INTEGRAL(RG_i - RU_i)$$

$$RG_i = \frac{R_{0i} - R_i}{RGT_i}$$

$$RGT_i = \frac{RGT_{0,i}}{TK_i}$$

$$RC_i = R_i/MRPR_i$$

$$MRPR_i = MRPR0_i + \frac{MRPRS_i}{\left(\frac{TK_i}{TK_{0i}}\right) * \left(\frac{C_i}{C_{i0}}\right) * \left(\frac{GG_i}{GG_{0i}}\right) * \left(\frac{GI_i}{GI_{0i}}\right)}$$

$$MRPRS_i = \frac{R_{0i}}{RU_{0,i}} * RA_{0,i} - MRPR0_i$$

$$RU = \min(RC, IRU)$$
$$IRU = RPI * RGRP$$

[17] https://www.gtap.agecon.purdue.edu/resources/download/6490.pdf

$MRPR_0$ = 4 Minimum exploitation Time (DOE/FOSSIL2)[18]
MRPRS = 6 (assumes lode life of 10 years (4+6))
RGRT = 1
AU= RU_0 =100
RPI = AU/RGRP as parameter initialization
RIT = 1.0 years (relevant economies dominated by agriculture)

AU = Average resource use
IRU = Indicated resource use
MRPR = Minimum reserve production ratio
MRPRS = Surplus (excess) MRPR
R = Resource
RA = Resource availability
RIT = Resource inventory time
RC = Resource capacity
RG = Resource generation
RGT = Regeneration time (10^{12} for now; as if all non-renewable)
RPI = Resource per industry
RU = Resource use

Violence

Violence is based on an increase relative to a status quo. The metric is the probability of experiencing violence, inferred from the population's violent behavior. Violence and conflict are considered synonymous. Conflict may have ideological rationalizations, but the primary motivation to participation is a choice relative to "legitimate" alternatives. Violence is not currently a function of immigrants.

For i=u,r,n

$$VP_i = 1/(1 + Exp(-VU_i))$$

$$VU_i = \alpha_i + \sum_{j=u,r,n} \omega_{ji} * Ln\left(\frac{AVP_j}{VP_{0i}}\right) + \omega_{4i} * Ln\left(\frac{WI_i}{WI_{0i}}\right) + \omega_{5i} * Ln\left(\frac{UER_{i,m}}{UER_{0i,m}}\right) + \omega_{6i} * Ln\left(\frac{GG_i}{GG_{0i}}\right)$$

For i=a,r
$VP_i = exo.f(time)$

[18] http://www.ntis.gov/search/product.aspx?ABBR=DOEPE7014302V2

AVP term is spillover violence from other regions.
VU = Violence utility

$\omega_{i,j}$ = 0.25 internal, i=j, 0.1 eternal i≠j
$\omega_{4,j}$ = -1.0 Pew study
$\omega_{5,j}$ = 1.0 Pew study
$\omega_{6,j}$ = 1.0 Pew study

Population and Migration

The population has three age groups with the middle group being the productive (p) work force. Migration decisions determine fractions of the population living in each location. Even moderately reliable migration data is only at the stock level; therefore the equations must include more complex logic to reflect the data available. Specifically, the stock level information allows the derivation of net-migration flows. Gross migration flows would be more advantageous. The use of an "exchange model" (ABM) representation would allow a better exploration of migration dynamics and an understanding of the differences between system dynamics (macro) versus agent-based (micro) simulation outcomes. The changes in population are due to births (BR), Deaths (DR), naturalization (NR), training (transition from common to skilled – TR), aging (transition to a different age group – AR), immigration (IR), and emigration (ER).

$$POP_{g,y} = Integral(BR_g - AR_{g,y} - DR_{g,y})$$

$$POP_{g,p} = Integral(AR_{g,y} - AR_{g,p} - DR_{g,p})$$

$$POP_{g,o} = Integral(AR_{g,p} - DR_{g,p})$$

$$BR_g = POP_{f,p} * BRATE_g$$

$$DR_{g,v} = POP_{g,v} * DRATE_{g,v}$$

BRATE and DRATE were calibrated to align population stocks with data.

To calculate fractions of the population in each location, we use a cognitive migration sub-model, based on a logit function as described in qualitative choice theory (McFadden 1974).

$$MP_{i,l,g,v} = Exp(MU_{i,l,g,v}) / \sum_k Exp(MU_{k,l,g,v})$$

$$MU_{i,l,g,v} = \alpha_{i,l,g,v} + \beta_{1i,l,g,v} * Ln\left(\frac{WI_i}{WI_{0i}}\right) + \beta_{2i,l,g,v} * Ln\left(\frac{GG_i}{GG_{0i}}\right) + \beta_{3i,l,g,v} * Ln\left(\frac{GI_i}{GI_{0i}}\right) + \beta_{4i,l,g,v} * Ln(VI_i) + \beta_{5i,l,g,v} * Ln(II_i) + \beta_{6i,l,g,v} * Ln\left(\frac{UER_{i,l}}{UER_{0i,l}}\right) + \beta_{7i,l,g,v} * Ln\left(\frac{POP_{i,x,l,g,m}}{POPx_{l,g,m}}\right)$$

"?I" terms are dissonance as an index (where"?" is a lead letter and "I" designates an index, as in VI). The rest are proportional cognition (instead of absolute differentials). Missing terms just come from further decomposing the α. Later determined, extraneous, terms just become constants that get subsumed in the α. It is assumed that none of the population generates any formal view of the expected future, other than as a consequence of a continuing trend. The dissonance representation is a "weak" (implicit) form" of expectation formation. The logarithmic (proportionality) construct for behaviors is better suited for extreme changes in conditions that the linear (normal/conventional BIA) construct.

$$II_i = \frac{WI_i}{AWA_i}$$
$$VI_i = \frac{VP_i}{AVP_i}$$

$$AWI_i = SMOOTH(WI_i, IAT)$$
$$AVP_i = SMOOTH(VP_i, VAT)$$

BRATE=Birth Rate
ARATE=Aging Rate
DRATE=Death Rate
MU=Migration Utility
DM=Disease Mortality
VP=Violence Prevalence
FA=Food Availability
WI=Wage Income
IAT=Income Averaging Time = 3 years
FAT=Food Averaging Time = 3 years
VAT=Violence Averaging Time = 3 years
AFA= Assimilated FA
AVP= Assimilated AVP
AWI= Assimilated WI
FI= Food Incongruity
VI= Violence Incongruity
II= Income Incongruity

Disease

This is a functional placeholder with the assumption that more detailed epidemiological model could replace this formulation.

$$DM_i = (\frac{CT_i}{CT_{0i}})^{\alpha_i} * (\frac{CP_i}{CP_{0i}})^{\beta_i}$$

α and β are different depending on the ecological zone.

α=1.1 Campbell-Lendrum 2003
β=0.55 Craig 2004, Thompson 2005

Food

This is a functional placeholder with the assumption that more a detailed agricultural model will replace this formulation.

$$FA_i = Min(1.0, \frac{FS_i}{FD_i})$$

$$FS_x = FP * RGRP_r * (\frac{CT_i}{CT_{0i}})^{\alpha_i} * (\frac{CP_i}{CP_{0i}})^{\beta_i}$$

$$FD_i = FPC * \sum_{i=u,r} \sum_{l,g,v} POP_{i,x,l,g,v}$$

FP = Food products per GDP
FS = Food Supply
FA = Food Availability
FD = Food Demand
FPC= Food per capital

α and β are different depending on the ecological zone.
α=-0.254 McCarl 2008
β=-0.041 McCarl 2008

FP:

0.304421	Neighbors
0.010356	USA
0.020388	ROW
	Mali
0.206555	Total

Water

This is a functional placeholder with the assumption that a more detailed hydrological model will replace this formulation.

$$WD_i = WPI_i * RGRP_i + WPC_i * \sum_{i=u,r} \sum_{l,g,v} POP_{i,x,l,g,v}$$

$$WA_i = Min(1.0, \frac{WS_i}{WD_i})$$

$$WS_i = WS_{0,i} * \left(\frac{CT_i}{CT_{oi}}\right)^{\alpha_i} * \left(\frac{CP_i}{CP_{oi}}\right)^{\beta_i} * \left(\frac{EE_i}{EE_{oi}}\right)^{\gamma_i} * WM_i^{\eta_i}$$

α=-1.3 Sheffield 2008
β=1.0 Default value until full hydro model in place.
γ=0.0 until definition of extreme events with normalization
η=0.0 for now as policy option

WPI and WPC are best obtained from SNL Hydrology dept. (Via WSM)
WPI=10, WPC=1, calculate WD0 using RGDP0 and POP0
Set WS0-1.5xWD for Mali and neighbors; 10 times for US and ROW

WS = Water Supply
WA = Water Availability
WD = Water Demand
WPC= Water per capital
WPC= Water per industrial activity
WM = Water Management (Scenario/policy)

Economic and Demographic Estimation

This uses a full times series from the World Banks plus UN on Real GDP and population by age for the initial time (year 1960 or greater) to year 2060, up to 2100.

For governance, we used WB wgidataset.xlsx and WI.xls, normalized. We exponentially interpolated between years, and assumed the last recorded value extended into the past. Governance for neighbors is weighted by WG GDP 2000$US.

LS (labor supply) is the population (15-64) by gender multiplied by labor participation rates. Total labor supply LS_T is the sum across gender.

Labor value (LV) is the RGDP times the labor fraction in EVFA.xls by skill type.

$$LV_{l,i,t} = RGDP_{i,t} * LF_{l,i}$$

Employment (assuming homogeneous unemployment across gender and skill), means employment by skill holds the rule:

$$\frac{E_{s,i}}{E_{u,i}} = \frac{LV_{s,i}}{LV_{u,i}} * WR_i$$

Where WR is the wage ratio of the main text.

Let: $\mu_i = \frac{LV_{s,i}}{LV_{u,i}} * WR_i$

Total Employment is:

$$E_{i,T} = \sum_{g} (LS_{i,s} + LS_{ic}) * (1 - UER_i)$$

Note this a sum for the "neighbor" countries.

Then the E for all time "t," for skilled (s) and common (c) is:

$$E_{i,s,t} = E_{T,i}/(1 + \mu)$$

$$E_{i,c,t} = E_{T,i}/(1 + 1/\mu)$$

EF is the employment fraction:

$$EF_{i,c,s} = \frac{E_{i,c,t}}{E_{i,c,t} + E_{i,s,t}}$$

$$EF_{i,s,s} = \frac{E_{i,s,t}}{E_{i,c,t} + E_{i,s,t}}$$

$$LW_{i,l} = LV_{i,l}/E_{i,l}$$

$$PGRP_i = PGRP_0 * \left(\frac{LS_{i,s}}{LS_{0i,s}}\right)^{LsF_i} * \left(\frac{LS_{i,c}}{LS_{0i,c}}\right)^{LcF_i} * \left(\frac{C_i}{C_{0i}}\right)^{CF_i} * \left(\frac{TK_i}{TK_{0i}}\right)^{(LcF_i + LsF_i)}$$

$$RGRP_i = PGRP_i * EC_i^{\alpha_i} * EG_i^{\beta_i} * EL_i^{\gamma_i} * RA_i^{\epsilon_i} * GG_i^{\mu_i} * GI_i^{\sigma_i}$$

Noted $EC_i^{\alpha_i} * EG_i^{\beta_i} * EL_i^{\gamma_i} * RA_i^{\epsilon_i}$ is definitely 1.0 for the referent case.

So:

$$RGRP_i = PGRP_0 * \left(\frac{LS_{i,s}}{LS_{0i,s}}\right)^{LsF_i} * \left(\frac{LS_{i,c}}{LS_{0i,c}}\right)^{LcF_i} * \left(\frac{C_i}{C_{0i}}\right)^{CF_i} * \left(\frac{TK_i}{TK_{0i}}\right)^{(LcF_i+LsF_i)} * GG_i^{\mu_i} * GI_i^{\sigma_i}$$

At time 0 all indices are unity except GG and GI, so:

$$RGRP_0 = PGRP_0 * GG_i^{\mu_i} * GI_i^{\sigma_i}$$

or

$$PGRP_0 = RGRP_0/(GG_0^{\mu_i} * GI_0^{\sigma_i})$$

$$C_0 = PGRP_0 * CF/ROI$$

The ROI is the rate of return (rent for capital) from GTAP. For all t:

$$PGRP_t = PGRP_0 * (RGRP_t/* RGRP_0) * (\frac{GG_t^{\mu_i}}{GG_0^{\mu_i}}) * (GI_t^{\sigma_i}/GI_0^{\sigma_i})$$

$$C_t = C_0 * PGRP_t/PGRP_0$$

$$IF_{i,t} = (C_{i,t} - C_{i,t-1} * \left(1 - \frac{1}{PL}\right))/RGDP_{i,t}$$

Note we have RGRP for all years including the future. We are not independently making the GRP forecast (not our expertise), but determining the impacts of climate and violence changes on that referent forecast. Also note that C is used to calculate PGRP using RGRP. PGRP reflects the “potential output of the economy” in the absence of detrimental government impacts.

Define:

$$\varphi_t = PGRP_0 * \left(\frac{LS_{i,s}}{LS_{0i,s}}\right)^{LsF_i} * \left(\frac{LS_{i,c}}{LS_{0i,c}}\right)^{LcF_i} * \left(\frac{C_i}{C_{0i}}\right)^{CF_i} * GG_i^{\mu_i} * GI_i^{\sigma_i}$$

$$TK_{i,t} = (RGRP_t/\varphi_{t_0})^{1/(LcF_i+LsF_i)}$$

The model now has the referent potential GDP, labor supply, employment, capital, wages, and technological advance for all time, gender, and region.

Migration Estimation

The OLS methods are simplest and can use the Berkson-Theil method if there are repeated observations,

However, the results contain biases. For example, given a multinomial logit (QCT) of the form:

$$P_i = \exp\left(\alpha_i + \sum_k B_k * X_{k,i}\right) / \sum_j \exp\left(\alpha_j + \sum_k B_k * X_{k,j}\right)$$

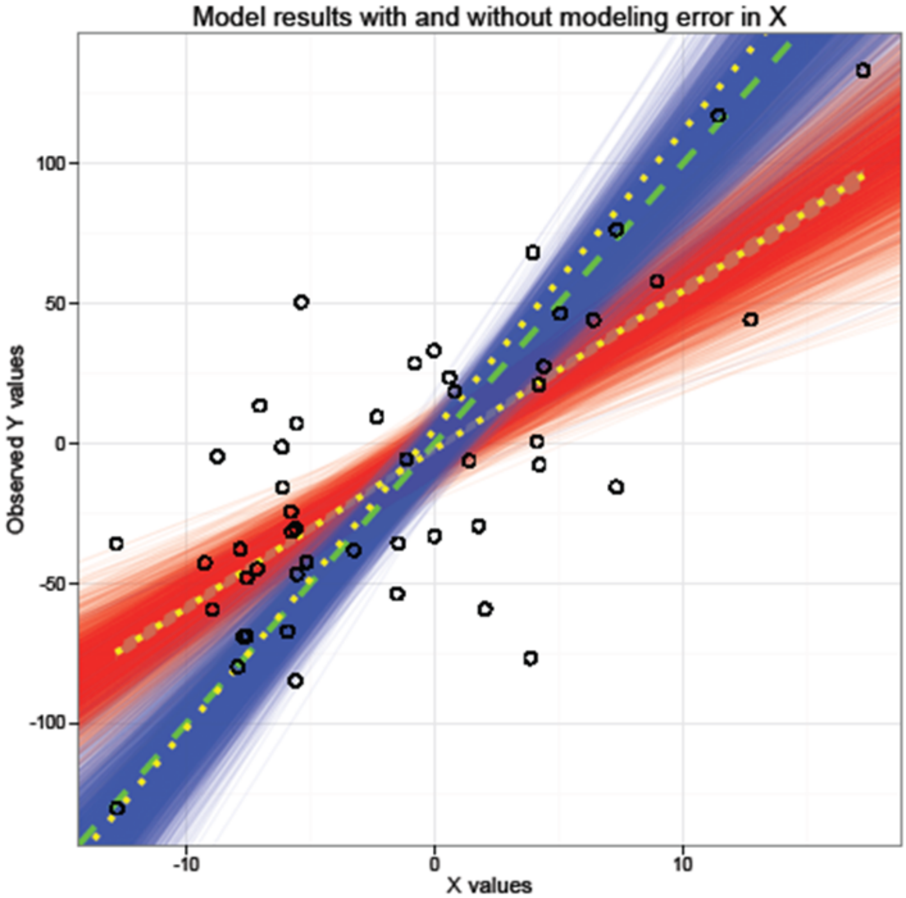
Where j is the set of all choices, for any particular observed choice "i" compared to a specific observed, alternative choice "J," and "X_k" are independent attributes affecting the choice:

$$Ln\left(\frac{P_i}{P_j}\right) = \left(\alpha_i - \alpha_j + \sum_k B_k * (X_{k,i} - X_{k,j})\right)$$

By setting one of the α as the numaire (usually as 0.0), the above equation is simply a linear-regression equation (Ben-Akiva and Lerman 1985).

Maximum likelihood methods are preferred in almost all instances having observational data and they are quite robust to a limited number of observations (Ben-Akiva and Lerman 1985).

The figure below shows the large difference between assuming the error is only associated with regression the shallow slope) and that obtained when uncertainty in the observations is recognized (Bayesian regression). The green line is the actual data generating process (DGP). The dotted lines are the "estimated values" from the resulting regression equation. The orange line within the red region is the regression assuming no observation uncertainty (definitional and asymptotically equals the dotted regression line). Both regressions were performed using maximum-likelihood.



Bayesian methods include direct Bayesian regressions that recognize the uncertainty in the observations. More sophisticated methods, typically using weighting (numerically transformed via ranking or scaling), allow the use of qualitative information form subject matter experts.

References:

Akiva, M. E. B., & Lerman, S. R. (1985). *Discrete choice analysis: theory and application to predict travel demand* (Vol. 9). The MIT press.

Backus, G. A., Lowry, T. S., & Warren, D. E. (2013). The near-term risk of climate uncertainty among the US states. *Climatic Change*, *116*(3-4), 495-522.

Campbell-Lendrum, D., A. Pruss-Ustun, and C. Corvalan. "How much disease could climate change cause." *AJ McMichael, D. Campbell-Lendrum, C. Corvalan, KL Ebi, AK Githeko, JS Scheraga, et al* (2003).

Craig, M. H., Kleinschmidt, I., Nawn, J. B., Le Sueur, D., & Sharp, B. L. (2004). Exploring 30 years of malaria case data in KwaZulu-Natal, South Africa: Part I. The impact of climatic factors. *Tropical Medicine & International Health*, *9*(12), 1247-1257.

Dunne, J. P., Stouffer, R. J., & John, J. G. (2013). Reductions in labour capacity from heat stress under climate warming. *Nature Climate Change*.

Fayissa, Bichaka, and Christian Nsiah. *The Impact of Governance on Economic Growth: Further Evidence for Africa*. No. 201012. 2010.

Kenefick, R. W., & Sawka, M. N. (2007). Hydration at the work site. *Journal of the American College of Nutrition*, *26*(suppl 5), 597S-603S.

McCarl, B. A., Villavicencio, X., & Wu, X. (2008). Climate change and future analysis: is stationarity dying?. *American Journal of Agricultural Economics*, *90*(5), 1241-1247.

Meadows, D. L., Behrens, W. W., Meadows, D. H., Naill, R. F., Randers, J., & Zahn, E. (1974). *Dynamics of growth in a finite world*. Cambridge, MA: Wright-Allen Press.

Miller, E. (2008). *An assessment of CES and Cobb-Douglas production functions*. Congressional Budget Office.

Sheffield, J., & Wood, E. F. (2008). Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, IPCC AR4 simulations. *Climate Dynamics*, *31*(1), 79-105.

Strauss, J. (1985). 10 The Impact of Improved Nutrition on Labor Productivity and Human-Resource Development: An Economic Perspective.

Thomson, M. C., Mason, S. J., Phindela, T., & Connor, S. J. (2005). Use of rainfall and sea surface temperature monitoring for malaria early warning in Botswana. *American Journal of Tropical Medicine and Hygiene*, *73*(1), 214-221.

Treyz, G. I. (1993). *Regional Economic Modeling. A Systematic Approach To Forecasting And Policy Analysis*. Kluwer, Boston

Van Art, Franz, Energy Efficiency in Power Plants, *ENERGY EFFICIENCY*. IN *IPPC INSTALLATIONS*. European Conference. Vienna, 21st and 22nd of October 2004, http://www.scribd.com/doc/85200470/6-Van-Aart

DISTRIBUTION

1	MS0899	Technical Library	9536 (electronic copy)
2	MS0359	D. Chavez, LDRD Office	1911

Sandia National Laboratories