

UNIVERSITY OF SOUTHERN CALIFORNIA

DOCTORAL THESIS

Forced Migration as a Cause and Consequence of Conflict

Author:

Cyrus Mohammadian

Supervisor:

Dr. Patrick James

*A thesis submitted in fulfilment of the requirements
for the degree of PhD.*

in the

Department of Political Science
&
School of International Relations

May 2016



Declaration of Authorship

I, Cyrus Mohammadian, declare that this thesis titled, 'Forced Migration as a Cause and Consequence of Conflict' and the work presented in it is my own. I confirm that this work submitted for assessment is my own and is expressed in my own words. Any uses made within it of the works of other authors in any form (e.g., ideas, equations, figures, text, tables, programs) are properly acknowledged at any point of their use. A list of the references employed is included.

Signed: 

Date: 7/12/16

“You have to understand, no one puts their children in a boat unless the water is safer than the land.”

Warsan Shire

Acknowledgements

I would like to express my sincere gratitude to my advisor, Dr. Patrick James, without whose guidance, patience, and expertise this dissertation would not have been made possible. I would also like to acknowledge my gratitude to the other members of my committee, Dr. Jefferey M. Sellers and Dr. John P. Wilson, for the assistance they provided at all levels of this research project. Finally, I would like to thank my dear friends and family, whose support and understanding have been vital to my success.

Contents

Declaration of Authorship	i
Abstract	iii
Acknowledgements	iii
Contents	iv
List of Figures	vi
List of Tables	vii
Abbreviations	viii
Symbols	ix
1 Introduction	1
2 Refugees and the Ethnic Balance of Power	4
2.1 Introduction	4
2.2 Demographic Pressure and Economic Competition	6
2.2.1 Refugees and Resource Scarcity	6
2.2.2 Refugees and the Ethnoreligious Balance of Power	8
2.3 Research Design, Data, and Methods	9
2.3.1 Research Design	9
2.3.2 Measurement and Data	10
2.3.3 Method	15
2.4 Results	15
2.4.1 Negative Binomial Regression	15
2.4.2 Robustness Check	17
2.5 Conclusion	18
3 When Ethnic Groups Rebel: Refugees and Transborder Kin	21
3.1 Introduction	22
3.2 Ethnic Armed Rebellion and One-sided Violence	22

3.3	Coethnic Refugees and Armed Ethnic Rebellion	24
3.3.1	Hypotheses	25
3.3.2	Methodological Approach and Data	25
3.3.3	Results	27
3.4	Coethnic Refugees and One-sided Violence	27
3.4.1	Hypotheses	29
3.4.2	Methodological Approach and Data	29
3.4.3	Results	31
3.5	Conclusion	35
4	The Logic of Population Control	37
4.1	Introduction	37
4.2	Literature Review	39
4.2.1	Ethnicity and Ethnic Civil Wars	39
4.2.2	Civilian Victimization	41
4.2.3	Forced Migration	41
4.3	The Logic of Population Control	42
4.3.1	Assumptions	43
4.3.2	Actors and Interests	45
4.3.3	Stylized Narrative	47
4.4	Hypotheses	48
4.5	Data and Methods	50
4.5.1	Methodological Approach	50
4.5.2	Estimation Technique	52
4.5.3	Data	54
4.6	Analysis	59
4.6.1	Organization of Results	59
4.6.2	Results	60
4.7	Conclusion	64
5	Conclusion	66
A	Appendix: Codebook	69
B	Appendix: Data Visualization	77
C	Appendix: Log Count Results	82
	Bibliography	87

List of Figures

2.1	Refugees as a Share of World Population since 1950	5
2.2	Refugees and Conflict April 2013 - May 2015	11
2.3	Conflict Events 2013-2015	12
2.4	Predicted Probabilities for High vs Low Tensions	17
4.1	Conflict States 1993-2009	54
5.1	Google Keyword Search Hits	68
B.1	Refugee Flows and Violence 2013-2015	77
B.2	Violence Severity 2013-2015	78
B.3	Violence by District 2013-2015	78
B.4	Syrian Refugees as a Share of Population 2013-2015	79
B.5	Ethnoreligious Tapestry of Lebanon	80
B.6	Ethnoreligious Groups by District	81

List of Tables

2.1	Conflict Events	16
2.2	Random Effects and Spatial Logit MCMC	19
3.1	Descriptive Statistics	27
3.2	Ethnic Refugees and Armed Rebellion	28
3.3	Descriptive Statistics	31
3.4	Ethnic Refugees and One-sided Violence	32
4.1	Raw Values of Selection Equation Variables	56
4.2	Binary Values of Selection Equation Variables	56
4.3	Log Count Values of Selection Equation Variables	56
4.4	Untransformed Variables in Selection Equation	57
4.5	Variables for Outcome Equation	58
4.6	Organization of Model Results	59
4.7	Marginal Effects Models	61
4.8	Conditional Effects Models	63
C.1	Organization of Model Results	82
C.2	Marginal Effects Models with Log Counts	83
C.3	Conditional Effects Models with Log Counts	84
C.4	Three-Stage (1) Heckit	85
C.5	Three-Stage (2) Heckit	86

Abbreviations

EBP	E thnic B alance P ower
CDC	C ategorically D isaggregated C onflict
ACD	A rmed C onflict D ata
EPR	E thnic P ower R elations
EAC	E thnic A rmed C onflict
ER	E thnic R efugees
RS	R esource S carsity
UNHCR	U nited N ations H igh C ommissioner for R efugees
UNRWA	U nited N ations R elief and W orks A gency
HAC	H eteroscedasticity and A utocorrelation C onsistent
UCDP	U ppsala C onflict D ata P rogram
WDI	W orld D evelopment I ndicators
GDP	G ross D omestic P roduct
IDP	I nternally D isplaced P erson
DV	D ependent V ariable
MLE	M aximum L ikelihood E stimation
PKK	P artiya K arkerên K urdistanê
ISIS	I slamic S tate of I raq and S yria

Symbols

T_l	probability of targeting loyal civilians
T_d	probability of targeting disloyal civilians
C	costs of targeting civilians
B	benefits of targeting civilians
V_e	level of violence against civilians during ethnic civil war
V_n	level of violence against civilians during nonethnic civil war
P_a	perceived probability of victimization in a civilian's potential destination of asylum
P_o	perceived probability of victimization in a civilian's country of origin
R	the number of refugees fleeing civil war
I	the number of IDPs fleeing civil war
F_e	share of forced migrants that are refugees fleeing ethnic conflict
F_n	share of forced migrants that are refugees fleeing nonethnic conflict
D_{it}	forced migration
Z_{it}	vector of explanatory variables
i, t	country, time (year)
u	error
f_{it}^*	ratio of refugee flows to total migrant flows in country i in time t
γ	vector of unknown parameters
ϕ	cumulative distribution function of the standard normal distribution
β	vector of known coefficients for explanatory variables
ρ	corr. of unobserved forced migrants and unobserved ratio of refugee flows to forced migrant flows
λ	inverse Mill's ratio
σ	standard deviation

*In dedication to my parents for nursing me with love and unconditional
support in my pursuit of success...*

Chapter 1

Introduction

Refugees in History

People have been fleeing war for as long as societies have waged it. It has been said that the Egyptian King Assyrophernes erected a massive monument in honor of his slain son and decreed that those who found shelter beneath it would be protected ([Phillipson 1911](#), 348). A similar right to asylum is related in the story of Assyrian King Ninus who erects a statue of his father that bestows protections to those seeking refuge from violence. Yet, it was the Israelites who first institutionalized the concept of a physical space of sanctuary, while the Saxon king of Kent, Ethelbert, was the first to formally codify it in law in the 6th century AD ([Rabben 2011](#), 59).

However, it was not until the late 15th century, during early modern Europe's emerging system of nation-states, that political authorities began to devote tremendous resources to population control. In contrast to the monarchies and religious authorities of previous eras, these leaders relied on their exercise of national sovereignty to legitimize their rule. The exercise of national sovereignty not only entailed establishing physical borders separating one nation-state from another, it also involved the delineation of sociocultural boundaries that ascribe and allocate national identity on the basis of various cultural markers, such as shared language and religious affiliation ([Marfleet 2007](#)). Central to this process was defining who the nation was and, as importantly, who it was not. Those whose national identities did not align with national fortunes of the state were either labeled as threats to the nation or were expelled -the "others".

The exodus of Muslims and Jews from the Iberian peninsula during the Inquisition, the flight of Puritans to the New World during the 16th and 17th centuries, and the migration of Calvinists of France and Germany to Protestant lands in the 17th and 18th centuries are prominent examples of these "others". In fact, it was the Huguenots, fleeing to England from a hostile Catholic France, who were first labeled *réfugiés*. It has been argued that the English monarchs, who

guaranteed their security, viewed them as “both an economic resource and an ideological asset” (Marfleet 2007, 140). As a result, these refugees became entangled in the process of statecraft. The French monarchs would remove Huguenots from the French body politic only for the English Kings to snub their counterparts by accepting them into theirs. This was hardly a feature unique to the early modern world. The Sassanians sheltered the Nestorian Christians, which outraged their Byzantine foes, and even the Greek city-states were said to have provided refuge for those defecting from their rivals. Tilly (1975) argued that states make war and wars make states. If he was right, then states also make refugees and refugees also make states. That is, states contribute to conditions conducive to the rise of refugees and the movement of refugees from one state to another contributes to the consolidation of the borders between them because giving sanctuary to one group also entails restricting it from others.

The purpose of this dissertation is to study forced migration as a cause and consequence of war-making and state-making in the modern era. I rely on statistical methods as well analyses of current events to show that states and rebels devote tremendous resources to the management of their populations -protecting some while driving others out. In the following section I preview the methods and key findings of each chapter.

Organization

Chapter 2 builds off recent quantitative work on civil war that has identified a link between refugee movements and the spread of conflict across borders. One commonly proposed mechanisms that accounts for this finding identifies refugee flows as a form of population pressure, which increases violence between host populations and incoming refugees. Another commonly proposed mechanism suggests refugees with ethnic ties to groups within the host population increase violence when their presence disrupts the ethnic balance of power between rival groups. Spatial regression results using a novel geo-coded dataset of substate violence in Lebanon between March 2013 - April 2015 reveal support for the latter mechanism.

Chapter 3 explores the ethnic balance of power theory in the context of two different forms of conflict -ethnic armed rebellion and one-sided violence -using global panel data on intrastate violence at the group and country levels. Although, the analysis fails to identify a significant association between the presence of coethnic refugees and ethnic armed conflict, coethnic refugees do significantly increase the probability of rebel led one-sided violence. I account for this by introducing a complimentary theory to EBP, “the logic of population control“, which contends that rebels have an incentive to target civilian supporters of rival groups as a means of undercutting the support these rivals enjoy. However, targeting civilians is costly if rebels cannot properly discriminate between their own supporters and civilians loyal to the government or

opposing groups. In ethnically polarized states, rebels can take advantage of the salience of ethnicity by mobilizing along ethnic lines and using ethnic markers to aid in the identification of potentially disloyal civilians. This logic, also applies to coethnic refugees. Rebels view refugees that share ethnic ties with the civilian populations represented by rival armed groups (government or otherwise) as a threat and respond to this threat with violence directed towards these unarmed civilians. Therefore, I conclude that while coethnic refugees do not increase the likelihood of conflict onset, they do increase the chances that civilians will be targeted, refugee or otherwise.

Chapter 4 examines why some civil war torn countries produce more refugees relative to their internally displaced population and others displace more of their population internally than across borders. Surprisingly, the relationship between internally displaced persons and conflict has been woefully underexplored. The aim of this chapter is to fill this gap in the literature. Using a panel dataset of civil conflicts by country-year from 1993-2010 and a two-step Heckman selection model, I show that civil wars fought along ethnic lines produce greater refugee flows relative to IDP flows than non-ethnic civil wars. I account for this finding by relying on insights drawn from the previous chapters. Specifically, I argue that in conflicts where combatants are recruited along ethnic lines, ethnic markers allow for less costly and more discriminate targeting of rival civilian populations, which in turn increases the share of forced migrants who seek refugee across borders relative to those displaced internally.

Chapter 5 concludes with a 1) review of this dissertation's key contributions, 2) a discussion of how these findings are situated within the larger literature, 3) and highlights a number of policy implications related to the issues of immigration and asylum policy.

Chapter 2

Refugees and the Ethnic Balance of Power

Abstract

Recent quantitative work on civil wars has identified a link between refugee movements and the spread of conflict across borders. One commonly proposed mechanism that accounts for this finding identifies refugee flows as a form of population pressure, which increases violence between host populations and incoming refugees. Another commonly proposed mechanism suggests refugees with ethnic ties to groups within the host population increase violence when their presence disrupts the ethnic balance of power between rival groups. Spatial regression results using a novel geo-coded dataset of substate violence in Lebanon between March 2013 - April 2015 reveal support for the latter mechanism.

2.1 Introduction

One of the most pressing humanitarian challenges facing the international community today is the plight of millions of refugees, asylum-seekers, and internally displaced peoples. In fact, the refugee crises is worse today than in any point in history. One of the most pressing humanitarian challenges facing the international community today is the plight of millions of refugees, asylum-seekers, and internally displaced peoples. In fact, the refugee crises is worse today than in any point in history. [Figure 2.1](#) shows the astonishing rise in the number of refugees and asylum-seekers as a proportion of world population since 1950. As of 2014, there are 54,937,556 refugees or individuals in refugee-like situations across 130 countries. Since 1995, the absolute number of worldwide refugees has increased by an alarming 170%. The majority of these people are

fleeing homes ravaged by violent conflict. Thus, much of the scholarly literature on the topic views refugees as a *consequence* of conflict and seeks to identify solutions to the crises (see for example [Vernant 1953](#); [Schechtman 1964](#); [Ferris 1993](#)). Yet, recent work in the field suggests refugees are a *cause* of war as much as they are a product of it.

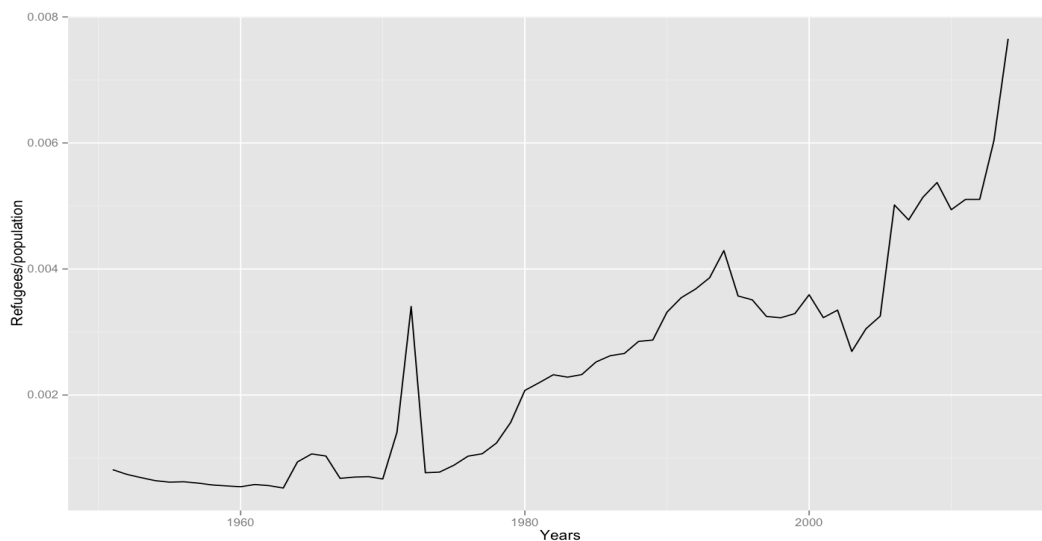


FIGURE 2.1: Refugees as a Share of World Population since 1950

For example, [Salehyan and Gleditsch \(2006\)](#), arguing that refugees contribute to conflict contagion, show that the presence of refugees from neighboring states increases the risk of civil war. [Bohnet \(2012\)](#) examines the relationship between refugee camp concentration and distribution and the outbreak of violence and finds that shorter distances between refugees and refugee settlements increases conflict risk. [Fisk \(2014\)](#) shows that refugee camps closer to the border tend to increase the risk of conflict.

However, most refugee flows do not result in conflict ([Lischer 2006](#)) and the extant mechanisms linking refugees to violence may tell us something about the variation in the geographic distribution of conflict *within* a state but does little to contribute to our knowledge regarding the variation in the geographic distribution of conflict between states. For example, the ongoing civil war in Syria has produced a hemorrhaging of refugees across most of its international borders. Our current models have correctly predicted that violence from the Syrian civil war is contagious yet they are unable to explain why that violence has spread in certain directions but not others. The mechanisms linking refugees to the spread of conflict remains under theorized.

The extant proposed mechanisms focus on the geographic distribution of refugee settlements at the substate level, such as distance between settlements, distance between settlements and major land markers such as capitals or international borders, and proximity to landscape favorable to rebellion. While these geographic factors are informative in identifying the location of conflict at the micro level, they cannot explain the variation in conflict likelihood across states at the macro

level. The problem is that these sorts of geographic variables tend to exhibit far more variation within a country than between countries. After all, the logical factors that dictate the location of refugee clusters within states, such as proximity to border crossings to ease the influx of refugees, are conditions that are likely faced by most states hosting refugees fleeing neighboring conflict. In contrast to geographic factors, however, demographic and economic factors can potentially explain variation both within and between states. As a result, mechanisms relying on variation in economic and demographic indicators should be identifiable at both micro and macro levels.

The focus of this chapter is to introduce two such mechanisms and test them at the micro level of analysis. I argue that the effect of refugees on conflict is conditional on two factors. First, refugees increase conflict levels by altering the ethnoreligious balance of power between rival groups. Second, refugees exacerbate competition over exceedingly scarce resources resulting in heightened levels of conflict. The following section discusses each mechanism in tandem.

2.2 Demographic Pressure and Economic Competition

2.2.1 Refugees and Resource Scarcity

The notion that population changes can lead to instability is not new; in his seminal treatise on population growth [Malthus \(1798\)](#) ominously predicted that human population growth would outpace agricultural production leading to a catastrophic breakdown in human socioeconomic structures. While his predictions failed to materialize, his ideas linking resource scarcity to conflict remained salient among a wave of postwar era scholars. These neo-Malthusianists argued that new technologies in mass agricultural production have led to a rapid increase in the global population, whose sustenance cannot be sustained with the world's increasingly depleted resources (see for example [Osborn 1948](#); [Vogt 1948](#); [Ehrlich 1968](#)).

Much of this work has since been decried as alarmist. What Ehrlich and others writing in his era failed to incorporate in their models is the phenomenon known as 'demographic transition', which entails a shift from high birth rates and death rates to low birth rates and death rates (without replacement) in first world countries ([Thompson 1929](#)). As a result the population growth of most developed countries has been outpaced by gains in productivity. While these theories failed to gain traction overtime, the idea that population shifts can lead to conflict continuous to spark scholarly attention.

More recently, scholars have begun to examine how short-term and rapid population changes can induce conflict over exceedingly scarce resources. In some instances, it has been argued that population growth and resource scarcity can actually work as a catalyst for technological

innovation (Boserup 1990). (Urdal 2005) finds no evidence linking population growth to conflict. On the other hand, Homer-Dixon and Blitt (1998) argue that developing countries are dependent on fresh water, cropland, forests, and fisheries –depletion of any of these resources they argue leads to conflict. Maxwell and Reuveny (2000) develop an agent-based model of population oscillations and warfare that largely follows the insights of Homer-Dixon and Blitt (1998). Raleigh and Urdal (2007) show that population growth and density are associated with conflict but not land degradation or water scarcity. In contrast, Theisen (2008) does find evidence linking land degradation and water scarcity with domestic conflict.

The general picture that emerges from these studies is that peace and stability are sensitive to population changes. Although a consensus on the mechanism underlying this relationship remains at large, the evidence does heavily weight in the favor of resource scarcity as a catalyst. Under conditions of resource abundance, rapid population growth can lead to the gradual depletion of these resources, which in turn can lead to conflict. Under conditions of resource scarcity, however, rapid population growth leads directly to conflict over the already scarce resources.

Borrowing from the literature on demography and conflict, one of the mechanisms this chapter tests is the role refugees play in bringing about demographic changes that lead to competition over scarce resources and ultimately violent conflict. For example, the war in Darfur between Arab northerners and non-Arab westerners resulted in a mass exodus of over 200,000 refugees into neighboring Chad beginning in 2003. The newly arrived refugees, most of whom were non-Arab Sudanese, were in immediate competition with locals over access to fresh water, government subsidies, firewood, etc.¹ By 2005 tensions had erupted between the refugees and the Arab population of Chad to the point that the non-Arab refugees began agitation movements alongside other non-Arab Chadians against the Arab Chadian majority; in less than a year Chad would collapse into an all out civil war as deadly as Sudan's (Reyna 2010).

Typically refugees fleeing to neighboring states are drawn from poorer segments in society, with little economic resources at their disposal, exposed to great dangers both environmental and man-made, and in great need of subsidized care, thus, placing them in direct competition with local populations (Loescher and Milner 2005). While international aid organizations are a great resource depended upon by many refugees, the greatest financial and political burden ultimately falls upon the host state. As a result, refugees place themselves in direct competition with other communities in the host state –coethnic or otherwise. Competition may arise over the use of arable land for crops, access to fresh water, employment opportunities, or even access to government services From this I derive my first two hypotheses:

H₁: *As the number of refugees a region hosts increases, so does its level of conflict.*

¹<http://tinyurl.com/j7da6w6>

H₂: *Resource poor regions will experience more conflict than regions rich in resources.*

2.2.2 Refugees and the Ethnoreligious Balance of Power

In contrast to the resource scarcity theory (RS), the ‘ethnoreligious balance of power’ (EBP) theory is couched in realist assumptions of conflict and cooperation. While realist theory of war originated in the field of international relations, many realist insights have been adopted in the study of domestic conflict (see for example [Posen 1993](#)). Peace between groups prevails when a certain balance of power between them emerges. When that balance of power is altered in one groups favor, however, the newly empowered group may attempt change to the way in which state rent is distributed in order to align it more closely with its own interests, which can result in conflict.

The sensitivity of the peace equilibrium to outside shocks is most pronounced under conditions of ethnic competition (*ibid*). According to Posen’s ethnic security dilemma, when states collapse and ethnic salience is a defining feature of political competition then conflict is likely to erupt due to the increasing inability to assure credible commitments and guarantees of peace between rival ethnic groups. In many circumstances, however, state collapse itself is endogenous to ethnic conflict. Nonetheless, Posen is right to suggest that exogenous shocks to an ethnically volatile system do in fact contribute to higher levels of conflict. Indeed state collapse does seem to induce ethnic conflict where ethnic tensions were previously kept in check. Similarly, it seems plausible to suggest that other exogenous shocks that introduce uncertainty or alter the balance of power between groups may also contribute to the onset of conflict, such as sponsorship and/or intervention by third parties or shifts in population numbers. Refugee flows are an example of the latter.

For instance, the Syrian civil war that erupted in 2011 lead to a hemorrhaging of Sunni refugees into Lebanon, undermines the fragile peace between Sunnis and Shias in the country that had emerged in recent years. Sunnis, whose numbers and political power along with Lebanese Christians have continued to dwindle relative to their Shia counterparts, have been emboldened by the presence of Sunni refugees fresh from the battlefields in Syria, leading to frequent clashes between Assad supporting Lebanese Shia (e.g. Hezbollah and Amal) and anti-Assad Sunnis. With a pool of fighters to draw from, the delicate ethnic balance between Sunnis and Shias in Lebanon has been jeopardized as Sunnis attempt to seize this rare opportunity to push back gains made by Hezbollah and their allies in Lebanon in recent years.

As illustrated in the case above, refugee flows can alter the ethnoreligious balance of power by providing ethnoreligious groups with a greater population base from which to extract resources and human capital (e.g. a greater potential pool of young fighters) and these flows can increase

the population concentration of those groups, which allows for more efficient mobilization. These flows modify the ethnic composition of the state when the refugees are composed of groups that share an ethnic identity with one or more constituent ethnic groups in the host state. Changes in the ethnic tapestry consequently lead to violence, particularly when relations between the competing ethnic groups are polarized. From this I derive my final two hypotheses:

H₃: *Regions with pre-existing ethnoreligious tensions are more likely to experience conflict.*

H₄: *The effect of refugees on conflict levels is magnified by the presence of pre-existing ethnoreligious tensions in a region.*

2.3 Research Design, Data, and Methods

2.3.1 Research Design

All extant micro level statistical studies of refugees and conflict have limited their empirical analysis to sub-Saharan Africa, which leaves us wondering whether the refugee conflict nexus is unique to sub-Saharan Africa or is generalizable to other regions of the world. This study extends the empirical scope of analysis to a previously unexamined region of the world, the Middle East. In particular, a substate analysis of the Syrian refugee crisis in Lebanon will be undertaken with the aim of testing the two theories discussed above.

Case selection is motivated by both theoretical and practical concerns. Lebanon is facing the most dire refugee crisis of all of Syria's neighbors -more than one in four residents in Lebanon is currently a registered Syrian refugee. Moreover, Lebanon and Syria share a similar composition of ethnoreligious groups, which is a necessary factor when testing for the conditional effects of sectarian tensions on refugee induced conflict spillover. Finally, Lebanon is the only state in the region to prohibit the establishment of formal UNHCR camps for its Syrian refugee population. With the notable exception of [Fisk \(2014\)](#), all previous micro level studies of refugees and conflict have limited their analysis to states with established camps. The effects of refugees dispersed among the host population remains largely unexamined.

Using Lebanon's twenty-six second order administrative districts as units of analysis, this study examines the spatial and temporal variation in the incidence of conflict for each month from March 2013 to April 2015. This results in a sample of 650 observations across 25 months ($N = 26$ districts x 25 months).

2.3.2 Measurement and Data

Conflict I begin with a discussion of my dependent variable. A novel conflict event dataset for Lebanon geo-referenced to the city level has been collected for this analysis. Over 2,500 conflicts have been recorded based on information gathered from two news wire services. For the purposes of this study, the data has been subset to include only the following event types:²

- Political executions
- Arrests and raids by military and internal security targeting militant groups
- Bombings (including suicide)
- Assaults and clashes involving militants and/or government (targeting civilians, government forces, or militant groups) as well as assaults and clashes arising from sectarian tensions (e.g. Lebanese civilian attacks on Syrian refugees)
- Protests
- Individual shootings linked to militants or linked to sectarian tensions (e.g. shootings between Alawi and Sunni families in Tripoli)
- Abductions by militant groups

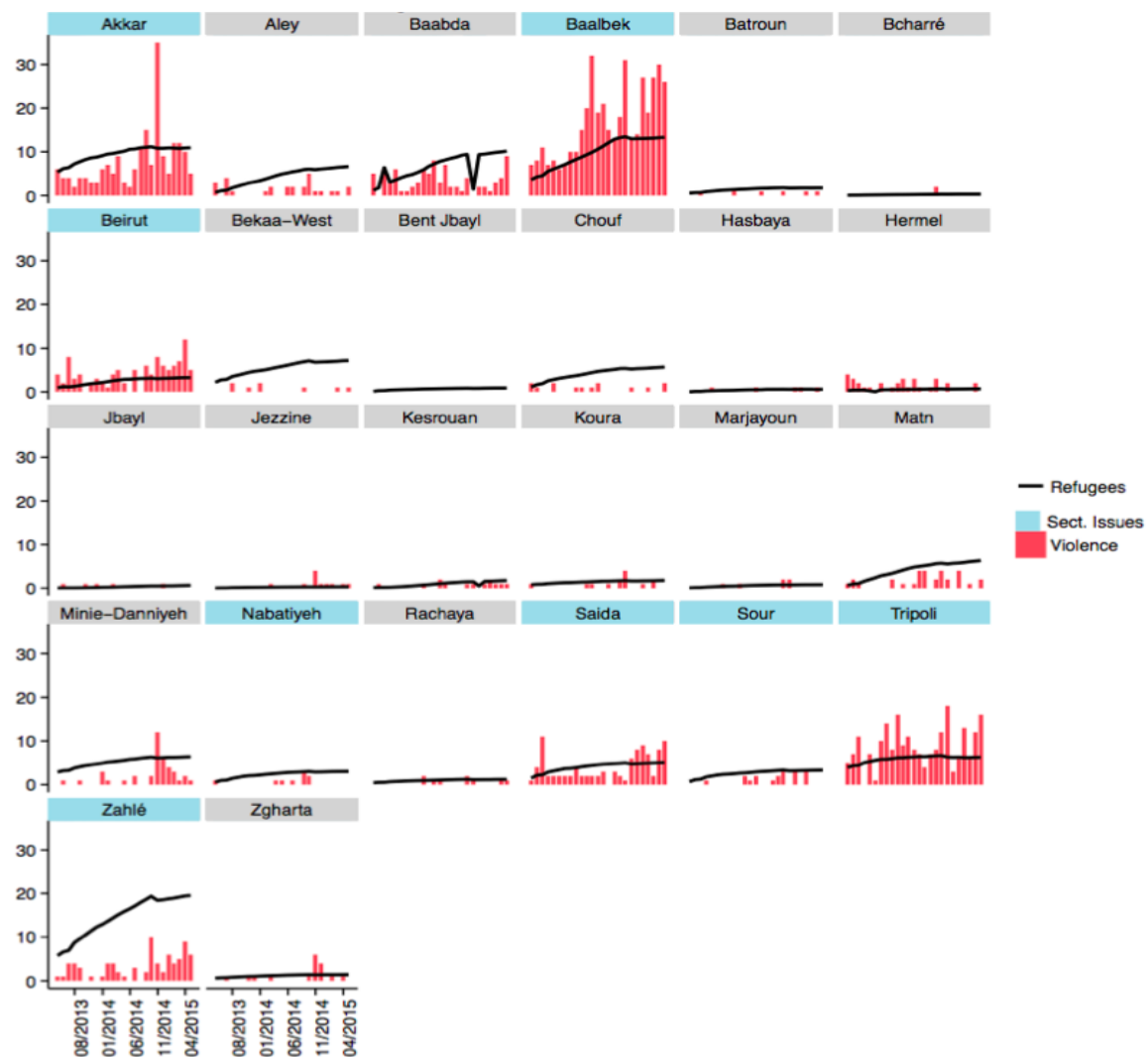
This result is a dataset of about 1,300 violent conflict events across 25 months across the entire country of Lebanon. [Figure 2.2](#) shows the number of violent conflict events over the entire sample period for each district in red. The number of violent conflict events has steadily increased for most districts and sharply increased for a few (Akkar and Baalbek) since data collection efforts began in 2013. [Figure 2.3](#) shows the spatial distribution of conflict over time. Conflict appears to cluster in space with concentrations in the northeast and central districts of the country.

Syrian Refugees Data on Syrian refugees in Lebanon is obtained from the [UNHCR](#) Syrian Refugee Response Team.³ Each month, the UNHCR releases map in PDF format of the total number of registered Syrian refugees for each district (not including Palestinians). Relying on this data, I construct a panel dataset with month-district as the unit of observation. Data at the district level was first made available in March of 2013, thus the temporal scope of this analysis begins in March of 2013 and extends to April of 2015 (the last date of the . Data on refugees for the month of August 2014 is missing; therefore, the missing data was generated using exponential interpolation. A time-series of refugee flows by district can be found in [Figure](#)

²Please refer to [Appendix A](#) for the dataset's codebook, which includes a more thorough discussion of the data collection procedure.

³<http://data.unhcr.org/syrianrefugees/regional.php>

FIGURE 2.2: Refugees and Conflict April 2013 - May 2015

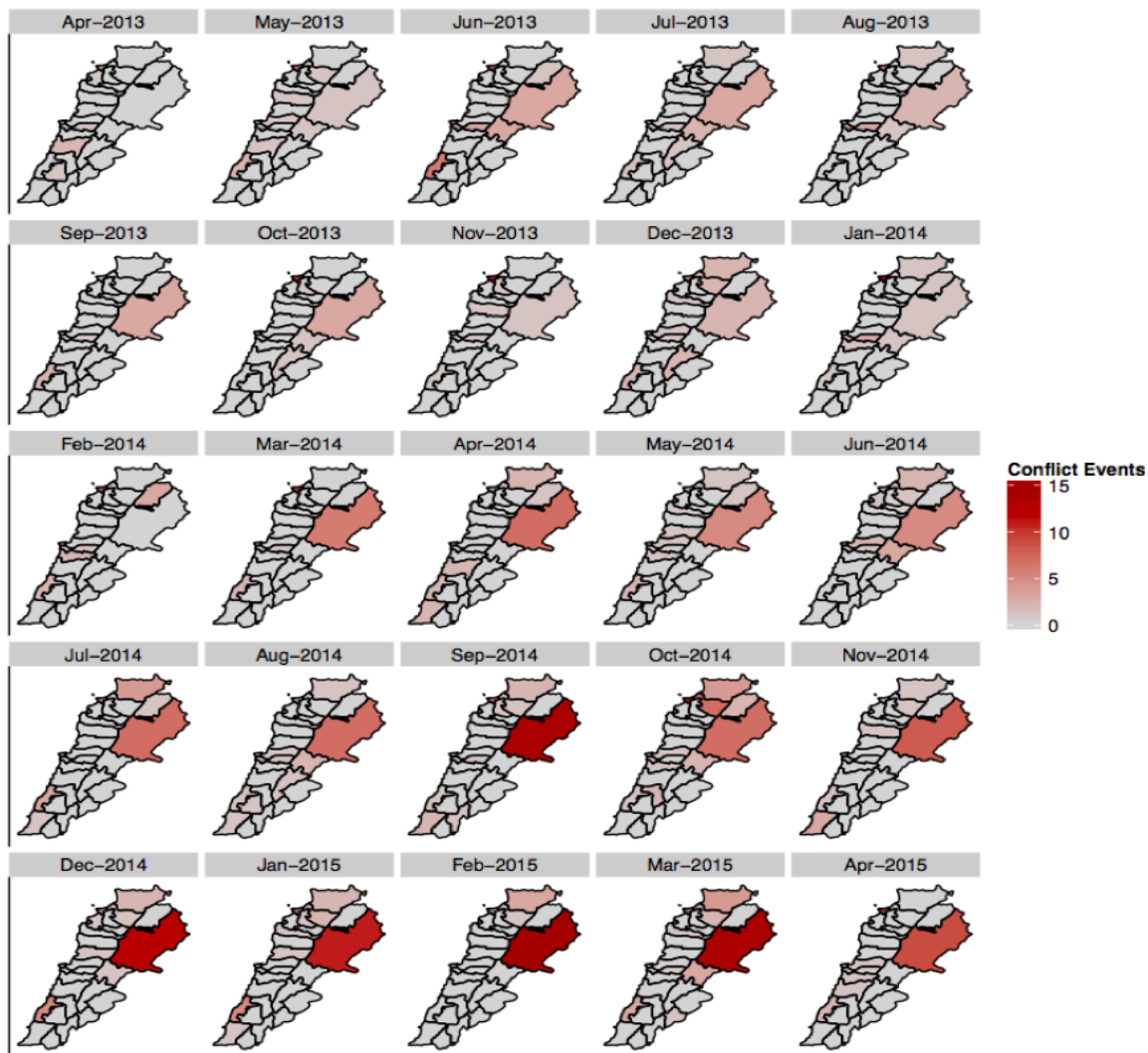


2.2. Following convention, I use the log of the number of refugees. It must be noted that these figures exclude Palestinian refugees fleeing Syria, who for the most part are relocated to existing Palestinian camps throughout the country. I expect higher refugee numbers to result in greater levels of conflict.

Palestinian Camps In addition to the massive influx of Syrian refugees, Lebanon also hosts a long-standing minority of Palestinian refugees who are located in twelve camps across eight different districts. Data on their geographic distribution is obtained from the [UNRWA](http://www.unrwa.org/palestinerefugees) for Palestinians in the Near East.⁴ This is a static and binary variable that does not capture changes to the number of Palestinian refugees over the sample period. I expect districts that host Palestinian refugee camps to experience more conflict than districts without camps.

⁴<http://www.unrwa.org/palestinerefugees>

FIGURE 2.3: Conflict Events 2013-2015



Population Density Districts facing greater levels of resource scarcity are expected to be more susceptible to conflict. One proxy for conflict used in the current analysis is population density. More densely populated districts impose greater strains on municipal governments, can drive down wages, and increase competition over access to government services. Therefore, I expect more densely populated districts to experience more conflict. Moreover, I test the conditional relationship of population density and refugees on conflict with the expectation that the positive impact of refugees on conflict levels is intensified in districts with higher levels of population density. I include a measure of population density (population/km²) based on Lebanese population records at the district level from 2007. District population varies from 21,650 people/km² (Beirut) to 62 people/km² (Jezzine), with a mean of about 1631 people/km².

Income Low income per capita has been shown to increase the likelihood of conflict at the country level (Fearon and Laitin 2003). However, the relationship between income levels and

conflict at the micro level is less clear. On the one hand, districts with lower income levels may face increased competition over access to government resources and employment opportunities, which can increase levels of conflict. On the other hand, conflict may correlate more with higher income districts than poorer ones because wealth can provide opportunities for looting (Hegre, Østby and Raleigh 2009). Thus, I include a measure of income that captures the percent of households in each district whose income is below the poverty threshold. Data on income levels is drawn from a national survey of households administered by Lebanon's [Central Administration of Statistics](#).⁵

Ethnoreligious Tensions According to the EBP theory introduced in this chapter, districts with a sensitive ethnoreligious balance of power are expected to suffer higher levels of conflict. Moreover, because I expect refugee flows to alter the delicate ethnoreligious balance of power between rival groups when refugees share a similar ethnoreligious make-up with the host population, I suspect the effect of refugees on conflict to be particularly pronounced in districts with sensitive ethnoreligious tensions. At its core, EBP theory involves two key factors, one related to characteristics of the host population and the other related to the characteristics of the incoming refugees. Not only must host populations contain rival ethnoreligious groups characterized by a delicate balance of power, the incoming refugees must include at least one group involved in this rivalrous relationship. Without both these factors in play, refugee flows would not alter the existing balance of power between these groups. Unfortunately, data on the precise ethnoreligious composition of incoming Syrian refugees is inaccessible. Moreover, Lebanon has not conducted a census since 1922, thus our knowledge of the distribution of ethnoreligious groups across Lebanon is also incomplete. Thus, in order to address these challenges, I make two assumptions. First, past studies have shown that refugees tend to flee to states where their co-ethnic kin reside (Moore and Shellman 2007; Rüegger and Bohnet 2015). In addition, the three largest ethnoreligious groups in Syria -the Alawi, Shia, and Sunnis -all have counterparts among the Lebanese population. Thus, I assume that among the refugees flowing from Syria into each of Lebanon's twenty-six districts, a substantial majority are either Sunni, Alawi, or Shia. Second, because the current primary political chasm in the country revolves around the divide between Shias and Alawis, on the one side, and Sunnis, on the other, I assume that the ethnoreligious balance of power in any district with substantial Sunnis and Shias/Alawis is at risk of shock due to the influx of Syrian refugees from across the border.

Based on these assumptions I operationalize pre-existing ethnoreligious tensions in two ways. First, any district with a mixed presence of Sunnis and Shia or Alawi is coded as having a sensitive ethnoreligious balance of power. Second, any two ethnoreligiously homogeneous districts that border one another are also coded as having a sensitive ethnoreligious balance of power if and only if the populations of the two districts in question are from rival groups

⁵<http://www.cas.gov.lb>

(Shia/Alawi vs. Sunni). The idea here is that although each district is homogeneous, their borders represent regions of the country where opposing groups are in close geographic contact with one another. Therefore, refugee influxes into these districts are also expected to increase levels of violence. As mentioned earlier, precise data on the distribution of ethnoreligious groups across has yet to be collected, therefore, I rely on approximate estimates based on of Lebanon's unique confessional voting system. Each ethnoreligious community has an allotted number of seats to represent their district in Parliament, which is suppose to be proportionate to each group's relative populations in that district. The threshold for representation per district is 5%. I use these seat ratios and population levels drawn from the year 2009 to reverse engineer the number of individuals from each ethnoreligious group for each district.⁶ Then, I drop any ethnoreligious group whose numbers fall below 10% of the total district's population. Finally, using the remaining ethnoreligious groups for each district, I follow the two coding procedures laid out above. In total, eight of Lebanon's twenty-six districts are coded as having pre-existing ethnoreligious tensions. These districts are identified by light blue highlights in [Figure 2.2](#).

Distance to Capital Previous work has demonstrated that conflicts tend to cluster in regions far from the capital ([Buhaug 2010](#)). As the distance from the capital increases, the projection of the central government's power diminishes and this affects the geography of conflict in two ways. First, armed groups opposing the central government are more likely to confront government forces in regions where the discrepancy between their capabilities is at its least. Second, the government's diminished power in regions far from the capital also hinder its ability to pacify intercommunal tensions. As a result, I expect higher levels of violence in districts further from the capital.

Neighboring Conflict As [Figure 2.2](#) shows, conflict seems to cluster in space. Thus, to account for this spatial clustering, I include a control for the levels of violence in all neighboring districts. I define neighbors as two districts that share at least one point in common. I expect levels of conflict in one district to be positively correlated with levels of conflict in neighboring districts.

I have developed a complimentary [website](#) and interactive web app using the [Shiny](#) platform.⁷ The web app allows the user to visualize data on conflict, population density, Syrian refugees, and the distribution of ethnoreligious groups in Lebanon using a series of choropleth maps, time series graphs, and bar charts. A series of hypervariate data visualizations displaying each of these variables can also be found in [Appendix B](#).

⁶This data is obtained from the [Lebanese Elections Data Analysis \(LEDA\) project](#).

⁷Note that the website takes some time to load: <https://cmohamma101.shinyapps.io/lebapp/>

2.3.3 Method

The research design and data discussed above result in four methodological challenges that must be addressed. First, the observations in the dataset likely exhibit temporal dependence. In other words, I expect conflict in one temporal period to increase the likelihood of conflict in subsequent temporal periods. Following Weidmann and Ward (2010), I address this problem in two ways. I employ robust standard errors (HAC) that control for temporal autocorrelation. Alternatively, I include in my model two temporal lags of the dependent variable. Second, because conflict events cluster in space, observations in my model are likely not independent of one another. Thus, in order to control for the effect of the presence of conflict events in neighboring districts I follow two different approaches. First, I cluster my standard errors by region to capture neighborhood effects. Alternatively, I include a first-order spatial lag of the dependent variable. A significant coefficient for the spatial lag suggests the level of conflict in one district is affected by the levels of conflict in adjacent districts.

Third, my research design faces potential problems of endogeneity. I address these concerns by lagging my time variant explanatory variables by one time period to ensure the causal arrow points from refugee flows to conflict and not the other way around. Finally, my dependent variable is a count measure that suffers from an overdispersion of zeroes. To account for this, I employ a negative binomial regression.

2.4 Results

2.4.1 Negative Binomial Regression

In this section I present the results of a series of regression models. I begin with an examination of my unconditional hypotheses. Model 1-3 employ robust standard errors clustered by administrative region (each of Lebanon's twenty-six second-order administrative districts are a member of one of six first-order administrative regions). Model 1 in Table 1 confirms H_1 , H_2 , and H_3 . Controlling for the presence of Palestinian refugee camps and proximity to the capital, Syrian refugee flows, high population density, and pre-existing ethnoreligious tensions all increase levels of violent conflict in Lebanon. While the presence of Palestinian refugee camps increases conflict levels as expected, the current analysis counterintuitively finds that conflict levels are actually higher closer to the capital.

Next I turn to the conditional hypotheses presented earlier. The analysis in Model 2 does not support the conditional hypotheses related to population density H_3 . Although population density increases conflict levels, its interaction with refugees is not significant at conventional

TABLE 2.1: Conflict Events

	(1)	(2)	(3)	(4)	(5)
Palest. Camps	0.557*** (0.126)	0.738*** (0.140)	0.659*** (0.123)	0.343** (0.139)	0.381*** (0.124)
Syrian Refugees	0.674*** (0.076)	0.463*** (0.090)	0.603*** (0.082)	0.481*** (0.088)	0.403*** (0.087)
Income	-0.001 (0.010)	-0.00004 (0.010)	0.001 (0.010)	-0.007 (0.010)	-0.010 (0.010)
Ethn. Tension	0.513*** (0.150)	-9.503*** (1.919)	0.571*** (0.153)	0.429** (0.178)	-4.877** (1.944)
Dist. to Cap.	0.012*** (0.005)	0.009* (0.005)	0.010** (0.005)	0.008 (0.005)	0.010** (0.005)
Pop. Dens.	0.0001*** (0.00001)	0.0001*** (0.00002)	-0.0004 (0.0003)	-0.0003 (0.0004)	0.0001*** (0.00002)
Ref. X Tension		0.925*** (0.172)			0.493*** (0.178)
Conflict.t1				0.061*** (0.012)	0.053*** (0.013)
Conflict.t2				0.041*** (0.013)	0.035*** (0.013)
Conflict.s1				0.018* (0.010)	0.024** (0.009)
Ref. X Pop.Dens.			0.00004 (0.00003)	0.00004 (0.00004)	
Constant	-7.645*** (0.875)	-5.559*** (0.980)	-6.967*** (0.923)	-5.597*** (0.959)	-4.851*** (0.922)
Observations	650	650	650	598	598
Log Likelihood	-996.315	-902.367	-993.897	-876.550	-873.983
AIC	2,006.630	1,820.733	2,003.794	1,775.100	1,769.965

Note:

*p<0.5; **p<0.01; ***p<0.001

levels and its effect size is near zero. [Model 3](#) tests the effect of refugees on conflict conditional on pre-existing ethnoreligious tensions. Controlling for all other factors, the effect of refugees on levels of violence is most acute in districts with pre-existing ethnoreligious tensions, which supports [H₄](#). This suggests EBP theory explains one mechanism by which refugee flows result in conflict.

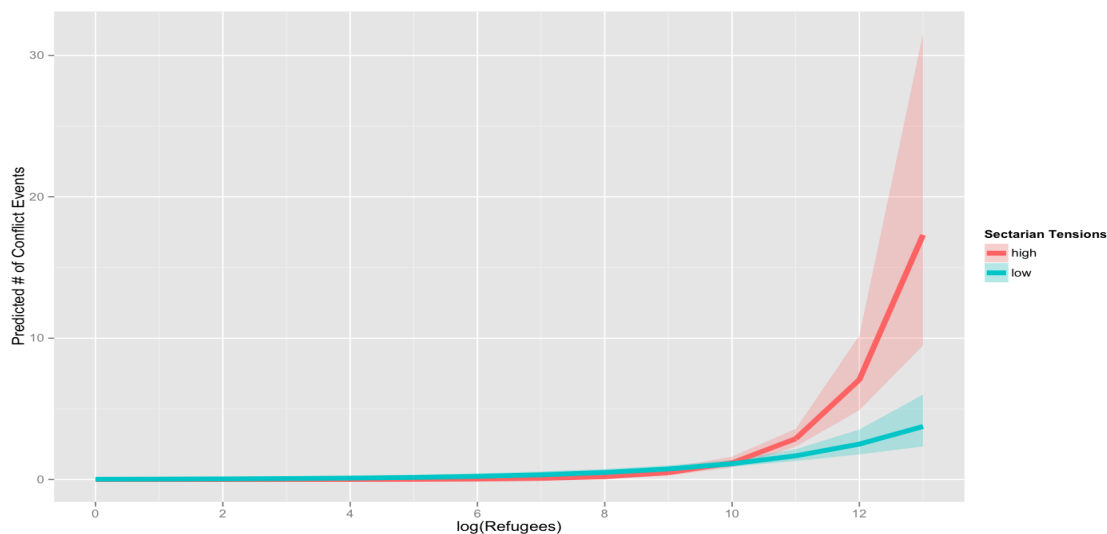


FIGURE 2.4: Predicted Probabilities for High vs Low Tensions

Models [4](#) and [5](#) introduce the temporal and spatial lags discussed earlier.⁸ The results from the previous models remain robust to the inclusion of spatial and temporal controls. The coefficient for both temporal lags is significant, suggesting that levels of conflict in one time period are influenced by levels of conflict in the two previous time periods. Moreover, the significant coefficient in front of the spatial lag shows that higher conflict levels in one district contributes to higher conflict levels in neighboring districts.

[Figure 2.4](#) shows the predicted probabilities (based on [Model 5](#)) along with confidence intervals for the effect of refugees on conflict for districts with and without pre-existing sectarian tensions. The difference between the effect of districts with and without ethnoreligious tensions on conflict clearly manifests itself at higher levels of refugee flows. In the following section, I re-estimate my model using random fixed-effects and novel spatial regression estimation technique that accounts for potential simultaneity bias in the research design.

2.4.2 Robustness Check

In this section, I implement two checks of robustness. First I rerun the model using random effects. Random effects are useful when data is organized hierarchically and observations are

⁸Robust standard errors are kept but are not clustering by region. Also note that including first and second-order temporal lags in the model drops the model's N to 598.

not independent of one another, typically this is the case with cross-sectional panel data. Observations are repeated at different time periods on the same units of analysis. Second, recent work suggests that adding spatial and temporal lags to conventional regression models results in simultaneity bias (Franzese, Hays and Schaffer 2010). Moreover, this endogeneity overestimates the coefficients of contagion variables at the expense of non-spatial variables. I employ a novel estimation technique using Markov Chain Monte Carlo developed by Chi and Zhu (2008) and implemented in the $\mathbb{R}$ computing language by Weidmann and Ward (2010). However, the implementation of this estimation technique has yet to be generalized to spatial count models. Thus, I run the analysis using a spatial autologistic model with a binary dependent variable. I code any district-month with more than two incidents of conflict as a 1 and all remaining district-months as a 0.

Random Effects The first two models in Table 2 shows the results of the RE estimation. Model 1 includes the first-order temporal lag while model 2 does not.⁹ Temporal lags have been dropped from the second model because temporal lags tend to correlate with the random intercept in RE models, which biases the coefficient of the lag variable (usually too large) and the coefficient of the explanatory variables (usually too small) (Ousey, Wilcox and Fisher 2011). As expected, the effect size on the conditional variable is reduced and drops in significance to the 0.10 threshold. However, in Model 2, where the temporal lag has been removed, the conditional variable's effect size increases and achieves significance at the conventional 0.05 level. The marginal effect of refugees on conflict in districts with no pre-existing ethnoreligious tensions also remains positive and significant, suggesting other mechanisms are also at work that tie refugees to the outbreak of violence.

Spatial Logit w/ MCMC The final model in Table 2 shows the results of the spatial logit model using the MCMC estimation techniques identified above. Once again, the conditional variable (Refugees X Ethnoreligious Tension) remains significant despite the transformation of the dependent variable from a count to a binary measure and the use of MCMC estimation techniques instead of Maximum Likelihood. The results of Model 3 suggest that the interaction of pre-existing ethnoreligious tensions and refugee flows increases the odds of conflict by a factor of 3.4.

2.5 Conclusion

Past studies have shown that refugees from neighboring countries increase conflict risk. In this study, I have investigated the mechanisms underlying that finding. Specifically, I introduced a

⁹The second-order lag is dropped from the model because its coefficient was insignificant and its respective model produces a higher AIC score suggesting that the inclusion of the second lag reduces overall model fit.

TABLE 2.2: Random Effects and Spatial Logit MCMC

	(1)	(2)	(3)
Palestinian Camps	0.555 (0.355)	0.612 (0.368)	0.909** (0.334)
Syrian Refugees	0.384** (0.139)	0.377** (0.0204)	0.576** (0.192)
Income	−0.0238 (0.0197)	−0.0248 (0.010)	−0.008 (0.021)
Ethnoreligious Tension	−3.94 (2.81)	−5.59* (2.72)	−12.721** (4.516)
Distance to Capital	0.0158 (0.00993)	0.0167 (0.0103)	0.014 (0.010)
Population Density	0.0000964* (0.0000404)	0.000104* (0.0000419)	0.0001 (0.0004)
CONFLICT.t1	0.0237* (0.0116)		0.778** (0.302)
CONFLICT.t2			0.713** (0.297)
CONFLICT.s1	0.0415*** (0.00799)	0.0441*** (0.0079)	0.038** (0.016)
Refugees X Tensions	0.436 (0.261)	0.595* (0.252)	1.225** (0.425)
Constant	−4.62** (1.57)	−6.967*** (0.923)	−8.769** (8.769)
Observations	624	650	598
Log Likelihood	−845.137	−847.777	1.075295
AIC	1716.3	1717.6	1173.855

Note:

*p<0.5; **p<0.01; ***p<0.001

model that explains transborder conflict contagion as a function of pre-existing ethnoreligious tensions that are exasperated by the arrival refugees of similar ethnoreligious stock. I posit that countries with polarized ethnoreligious configurations are more susceptible to conflict contagion. When transborder refugees share similar ethnoreligious configurations with the host population they can alter the balance of power between rivals groups, resulting in higher levels of conflict. In addition, I argue that this mechanism should be identifiable at macro and micro levels because substantial variation between units is expect at both levels. I test this theory against an alternative that identifies refugees as a form of demographic pressure that afflict regions suffering from resource scarcity most acutely.

The results of my analysis confirm past studies that have identified a significant and positive relationship between refugees and conflict. Moreover, I find support for the notion the effect of refugees on conflict is conditioned by the ethnoreligious configurations shared by the refugees and host population. Regions characterized by delicate ethnoreligious configurations are at particular risk for conflict contagion. Mechanisms linking the interaction of refugees and resource scarcity to conflict, however, are unsupported by the data. While the marginal effect of population density on conflict is positive and significant, low income is insignificant as is the conditional effect of refugees and population density on conflict levels. These findings are robust to a number of different specifications and estimation techniques.

What remains unclear is what type of violence the presence of refugees is most associated. Do refugees increase violence among militants? Between government forces and rebels? Or does the presence of refugees increase the chances these groups target noncombatants? In the following [chapter](#), I test the ethnic balance of power theory in the context of two different forms of violence. First, I examine coethnic refugees and ethnic rebellion using a global dataset of disempowered ethnic groups from 1975-2009. Second, I examine the relationship between coethnic refugees and one-sided violence committed by government and rebel groups using a global country-year dataset. The results suggest that coethnic refugees have no significant effect on the likelihood of ethnic rebellion but do increase the chances that rebels will target noncombatants when they do fight. I account for this finding by introducing a complimentary theory to EBP, the logic of population control, which argues warring parties face incentives to control populations in contested territory by protecting loyal civilians and targeting those sympathetic to rivals.

Chapter 3

When Ethnic Groups Rebel: Refugees and Transborder Kin

Abstract

In this chapter, I explore the ethnic balance of power theory in the context of two different forms of conflict -ethnic armed rebellion and one-sided violence -using global panel data on intrastate violence at the group and country levels. Although, the analysis fails to identify a significant association between the presence of coethnic refugees and ethnic armed conflict, coethnic refugees do significantly increase the probability of rebel led one-sided violence. I account for this by introducing a complimentary theory to EBP, "the logic of population control", which contends that rebels have an incentive to target civilian supporters of rival groups as a means of undercutting the support these rivals enjoy. However, targeting civilians is costly if rebels cannot properly discriminate between their own supporters and civilians loyal to the government or opposing groups. In ethnically polarized states, rebels can take advantage of the salience of ethnicity by mobilizing along ethnic lines and using ethnic markers to aid in the identification of potentially disloyal civilians. This logic, also applies to coethnic refugees. Rebels view refugees that share ethnic ties with the civilian populations represented by rival armed groups (government or otherwise) as a threat and respond to this threat with violence directed towards these unarmed civilians. Therefore, I conclude that while coethnic refugees do not increase the likelihood of conflict onset, they do increase the chances that civilians will be targeted, refugee or otherwise.

3.1 Introduction

Recent literature in the field has identified a link between refugees and the spread of conflict across borders. One causal mechanism that I have offered to account for the findings of previous work is the role that shared ethnic ties play as a catalyst for conflict. I argue that refugees that share ethnic ties with politically active groups in their host state can alter the existing ethnic balance of power (EBP) between competing groups. If tensions are hostile between these rivals, then even minor shifts in demographics that advantage one group over another can increase the number of violent events.

In the previous chapter, I tested this proposition at the substate level in Lebanon and found evidence to support it. In this chapter, I test this EBP theory globally at the country and group levels in the context of two different forms of conflict -ethnic armed rebellion and one-sided violence. Although, the analysis fails to identify a significant association between the presence of coethnic refugees and ethnic armed conflict, coethnic refugees do increase the probability rebels will target unarmed civilians. I account for this finding by introducing a complimentary theory to EBP, "the logic of population control", which contends that governments have an incentive to target civilian supporters of rival groups as a means of undercutting the support their rivals enjoy. However, targeting civilians is costly if governments cannot easily discriminate between their own supporters and the civilians loyal to the opposition. The conditions are different when rival groups mobilize along ethnic lines because ethnic markers aid in the identification of potentially disloyal civilians. This logic, also applies to refugees who share ethnic ties with these same groups.

3.2 Ethnic Armed Rebellion and One-sided Violence

While the empirical approach of the previous chapter leveraged its disaggregated unit of analysis to uncover interesting links between demographic shifts and outbreaks of violence, the approach adopted suffers from three data limitations that can be overcome by moving the unit of analysis either to the country or ethnic group level.

First, due to data limitations, the analysis of the previous chapter took for granted the ethnic background of the Syrian refugees flowing into Lebanon. The data on Syrian refugees at the substate level in Lebanon that the UNHCR has released does not include demographic data, such as religious or ethnic affiliation. As a result, I have approached the empirical analysis with the assumption that Syrian refugees tend to migrate to regions of Lebanon inhabited by Lebanese whom they share ethnic and/or religious ties with. If this assumption holds true, however, I face another challenge. If ethnic refugees are the only form of refugees in a particular

region then I cannot know if it is their shared ethnic status that is responsible for the increased chances of conflict or if it is simply because they are forced migrants who bring with them weapons and dangerous ideologies from the battles fields they fled. Recent released data on the ethnoreligious background of refugees (ER dataset) can help ameliorate this problem ([Wimmer, Cederman and Min 2009](#)). Because this data is not territorially disaggregated, a novel country or group year analysis must be undertaken.

Second, the logic behind the theory of EBP posits that coethnic refugees are a potential source of recruitment for combatant groups. Therefore, we should expect groups who enjoy the benefits of recruiting from larger pools of coethnic refugees to be more likely to rebel against the state. Although the substate data collected for the empirical investigation of the previous chapter includes demographic information on the victim and perpetrator of each conflict event, records for all observations are incomplete and analogous demographic information pertaining to refugees at the substate level does not exist. The ACD2EPR dataset includes group level data on the outbreak of ethnic civil war ([Vogt et al. 2015](#)), which combined with the ER dataset ([Rüegger and Bohnet 2015](#)), can form the basic design of a group-year analysis that examines the relationship between the presence of coethnic refugees and the likelihood of rebellion for each country's ethnic groups.

As mentioned above, while the data from the previous chapter also includes information on the identity of victims and the perpetrators of violence, that sort of information is simply unavailable for many of the observations in the dataset. Therefore, while the analysis of the previous chapter revealed a statistically significant link between refugees and violence, it is unable to identify what type of violence the presence of refugees is increasing. This is problematic for our understanding of the mechanics behind the refugee-conflict nexus. Without this information, we cannot know, for example, if the presence of refugees is destabilizing because these refugees are actively involved with fighting or if the increase in violence is a result of hostile groups targeting civilian refugees. Moving the analysis to the country level allows me to leverage data on one-sided violence released by the Uppsala Conflict Data Program (UCDP) to test this proposition ([Eck and Hultman 2007](#)).

Finally, due to the difficulty in obtaining accurate geocoded data at the substate level, the analysis of the previous chapter was confined to the state of Lebanon. Using data at the group and country level allows me to expand the geographic scope of this analysis beyond the Levant. Thus, the purpose of this chapter is to provide a further test of EBP theory at two new levels of analysis, across new regions of the world, and with data that allows me to test all parameters of my theory at once. To that end, in this chapter I will introduce a group-year analysis that examines the impact of refugees on the likelihood of rebellion for their coethnic brethren in their

host states. I will also present the results of a separate country-year analysis that looks at the impact of coethnic refugees on one-sided violence in host states.

Ultimately, what I find is that refugees exert no statistically significant impact on the likelihood that their coethnic groups in the host state will rebel, but they do increase the probability of rebel led one-sided violence. On the one hand, both the substate and the country-year analyses revealed a substantial link between refugees and conflict conditioned by the presence of volatile ethnic tensions. On the other hand, the group-year analysis fails to uncover a significant relationship between ethnic conflict onset and the presence of ethnic refugees. I account for this discrepancy by introducing a complimentary theory to EBP, which I label the “logic of population control”. I argue that the presence of coethnic refugees can alter the ethnic balance of power, but changes in the ethnic balance of power between rival groups contributes more to how conflict is fought than to the specific likelihood of its occurrence. In other words, changes in relative power between rival ethnic groups impact the tactics of violence adopted by warring parties. The presence of refugees with ethnic ties to one group, I argue, increases the incentives rival groups have to target refugees and their local coethnic civilians.

The logic of population control contends that refugees are a valuable source of human capital in the form of recruits, public solidarity, and even economic activity. Rival groups are keenly aware of this and, thus, rightly or wrongly, view these refugees as a threat that needs to be controlled. Thus, coethnic refugee flows to states with volatile ethnic relations increases the likelihood of civilian targeted one-sided violence as rebels (or governments) target refugees with the aim of undercutting their support for their coethnic brethren.

The remainder of this chapter is divided into two sections. In the first section, I review the design and results for the analysis of the effect of refugees on the probability of ethnic group rebellion. In this second section, I introduce the design and present the results of the country-year analysis that reveals a link between the spread of coethnic refugees across borders and one-sided violence. I conclude with a discussion of a number of important observable implications of this theory, which I proceed to test in the following [chapter](#).

3.3 Coethnic Refugees and Armed Ethnic Rebellion

In the previous [chapter](#) I examined how Syrian refugees contributed to increases in the levels of violence at the substate level in Lebanon. I found that refugees do increase conflict but I also discovered that the impact was even stronger for districts with previously volatile ethnic relations. Unfortunately, due to the limitations of the data collection process, I was unable to identify precisely what kind of violence this entailed. Thus, it is unclear whether this increased

conflict takes the form of violence between rival armed groups emboldened by the arrival of pools of new recruits or if it is in the form of violence targeted specifically at refugees and other unarmed civilians. Put more simply, are refugees actively causing this violence as much of the media and even some academic work has framed the topic or are they victims of violence in the very places they have sought refuge? I argue that the empirical analysis of this chapter lends support to this latter claim. In this section, I have designed a ethnic group-year analysis that leverages newly released data on the ethnoreligious background of all refugees between 1975-2009, which I use to test the proposition that the presence of refugees emboldens their coethnics in the host state to rebel.

3.3.1 Hypotheses

Given the findings of the previous chapter, I argue that refugees increase the probability of conflict by altering the demographic balance between rival groups. Because they are a potential source of recruitment for groups they share ethnic ties with, rivals of these groups (or rivals in government) view these refugees as a potential threat. This destabilization is what EPB theory argues contributes to conflict in the first place. The presence of refugees either emboldens their coethnic brethren in the host state, agitates rivals of these groups, or both; the result in each scenario is more conflict. From this, I derive my primary hypothesis:

H₁: *The presence of coethnic refugees emboldens ethnic rivals excluded from sharing power with the government to rebel.*

3.3.2 Methodological Approach and Data

The methodological approach adopted in this section is designed to examine whether the presence of refugees with ties to host ethnic groups increases the likelihood these groups rebel against the government. Thus, I use a logistic regression model with country-group-year as my unit of analysis. That is, each observation is a record of each politically active ethnic group in each country in each year between 1975-2009. The dependent variable in this model is binary and indicates the onset of rebellion by a specific ethnic group in a particular country. I employ robust standard errors clustered by country-group (some ethnic groups exist in more than one country so I cluster using within country ethnic groups) to account for the nested nature of the panel data and I introduce region-fixed effects to control for unobserved heterogeneity from one region of the world to another. I also include a count for the number of years of peace since the last rebellion as well as its squared and cubed term to account for temporal autocorrelation *a la* [Carter and Signorino \(2010\)](#).

The data used for the dependent variable, armed rebellion by an ethnic group, is a recoding of UCDP/PRIO's Armed Conflict Data (ACD) by the coders of the Ethnic Power Relations (EPR) dataset at the University of California, Los Angeles (Wimmer, Cederman and Min 2009). The result is a dataset (the Ethnic Armed Conflict (EAC) dataset) of all politically active ethnic groups across the world categorized according to the "degree of access to central state power by those who claimed to represent them" (*ibid*, 326). It also includes the dependent variable used in this study –a binary indicator of the onset of armed ethnic rebellion. Armed ethnic rebellion is defined as armed conflict where at least 25 combatants have been killed annually and "at least one party is the government of a state" (*ibid*, 326). What makes it "ethnic" is that the rival party must be an armed group that makes "ethnic aims" (such as achieving ethno-national autonomy) and the armed rebels must be mobilized along ethnic lines (i.e. recruits are drawn primarily from the same ethnic group and alliances are forged along ethnic ties). Once an ethnic group rebels, years of ongoing conflict are dropped from the analysis because this is a model that is testing the onset of ethnic conflict and not its duration. Furthermore, only ethnic groups excluded from power are included in the analysis. Therefore, if the ruling government is composed of parties representing one or more ethnic groups, those ethnic groups are not included in the analysis because by definition they cannot rebel against themselves. It must be noted that requiring one party to the conflict to be a government of a state, omits a wide range of forms of violence, such as violence against noncombatants, violence between rebels, terrorism, and even genocide.

The primary independent variable of this study is the presence of coethnic refugees. A recently released dataset of the ethnic background of all refugee flows from 1975-2009 has been released, known as the Ethnicity of Refugees (ER) dataset (Rüegger and Bohnet 2015). This dataset takes the dyadic data on refugee flows between countries released by the UNHCR and identifies up to three different ethnic groups per flow –a majority group and up to two minority groups using information drawn from experts as well as primary sources. Each group is allotted a particular share of the total number of refugees of each year and aggregate numbers are then reverse engineered from these proportions. I log the total value.

I include a host of group and country level variables in the analysis. In terms of group level variables, I include, the log of a group's population, the log of a group's population as a share of total country population, the history of conflict between the ethnic group and the state (operationalized as the total number of ethnic conflicts between the group and the government of that state), and whether the ethnic group had lost a share of power in the ruling coalition in the previous two years. In terms of country level controls, I include whether the state is a democracy, which I obtain from the Polity IV dataset (Marshall and Gurr 2014) and I use the log of GDP/capita drawn from the World Bank's World Development Indicators (WDI). I also control for unobserved region-to-region heterogeneity using region fixed-effects. All variables

are lagged by one year. Table 3.1 shows the descriptive statistics for each of these variables. Next, I turn my attention to the results.

TABLE 3.1: Descriptive Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Ethnic Conflict	251,580	0.005	0.067	0	1
Democracy	247,937	0.202	0.402	0	1
Previous Conflict	251,580	0.076	0.297	0	3
log(Share of Population)	251,580	1.008	1.098	0.002	4.595
log(Ethnic Refugees)	251,580	0.074	0.886	0.000	14.950
log(GDP/capita)	250,978	8.266	1.108	5.179	15.010
log(Population)	250,978	11.715	1.863	6.272	14.094
Peace Years	251,506	38.721	15.493	0	63

3.3.3 Results

Table 3.2 shows the results of the logistic regression model with robust standard errors in parentheses. All three peace year terms are significant, indicating a clear time trend in the outbreak of rebellion over the sample time period. Of the remaining controls, previous conflict history and recent downgrade in power sharing status increase the chances of ethnic group rebellion, while armed groups representing ethnic groups with greater populations are actually less likely to rebel against the government. Most importantly, H_1 is unsupported by the data, which suggests there is not enough evidence to support the claim that refugees increase the chances of armed rebellion by their transborder kin. This says nothing, however, of levels of one-sided violence that may emerge in response to the presence of refugees. In the next section I examine whether the presence of refugees with shared ethnic ties can contribute to increased chances of civilian targeted violence.

3.4 Coethnic Refugees and One-sided Violence

The previous section failed to uncover evidence linking the presence of coethnic refugees with more ethnic armed conflict. In this section, I examine whether refugees alter the ethnic balance of power in ways that increase one-sided violence against civilians.

TABLE 3.2: Ethnic Refugees and Armed Rebellion

	<i>Dependent variable:</i>
	Onset of Ethnic Conflict
Intercept	−0.33 (2.43)
Democracy	0.18 (0.39)
log(Ethnic Refugees)	−0.75 (1.01)
log(Share of Pop)	0.23 (0.12)
Previous Conflict	1.18** (0.18)
log(GDP/capita)	−0.46* (0.17)
log(Population)	−0.24* (0.08)
Peace years	−0.17** (0.05)
Peace years ²	0.01** (0.00)
Peace years ³	−0.00** (0.00)
Downgraded Status	2.03** (0.37)
Asia	1.25 (1.13)
Eastern Europe	1.33 (1.10)
Sub-Saharan Africa	0.66 (1.16)
North Africa/Middle East	1.43 (1.12)
Latin America	0.01 (1.20)
Observations	247,865
AIC	12803
χ^2	1,694.988*** (df = 15)
<i>Note:</i>	*p<0.05; **p<0.01

3.4.1 Hypotheses

Although the previous section did not uncover a statistically significant relationship between ethnic armed rebellion and the presence of coethnic refugees, coethnic refugees may impact other forms of conflict, such as violence aimed at unarmed civilians. In fact, I posit that the presence of refugees with ethnic ties to the host population *does* alter the ethnic balance of power between groups but that these changes do *not* make armed conflict more likely. Instead, they *change* the way in which armed conflict is waged. Specifically, I argue that refugees are a valuable source of human capital in the form of recruits, public solidarity, and even economic activity. Rival groups are keenly aware of this and, thus, rightly or wrongly, view these refugees as threats to be controlled. Thus, the presence of coethnic refugees increases the likelihood of one-sided violence as rebels (or government forces) attempt to control the population by targeting refugees with the aim of undercutting the support they provide to their coethnic brethren in the host state.

I label this the logic of population control because warring parties have an interest in controlling populations sympathetic to their own cause while undercutting the support of populations their rivals enjoy. Because coethnic refugees are identifiable by the same ethnic markers that place their local coethnic brethren at risk of targeting, they too are at risk of persecution. Governments as well as warring rebels have an interest in targeting the civilian support base of their rival groups. However, targeting civilians is costly if conflicting parties cannot distinguish between sympathizers and those sympathetic to their rivals. If these conflicting parties can identify supporters based on ethnic markers, this task is eased. Refugees, like the local civilian population, then are at risk of persecution if they share ethnic traits with these locals. From this, I derive my remaining two hypotheses.

H₂: *Rebels in states hosting coethnic refugees are more likely to target civilians during conflict than rebels in states that do not host coethnic refugees.*

H₃: *Governments of states hosting coethnic refugees are more likely to target civilians during conflict than government's of states without coethnic refugees.*

3.4.2 Methodological Approach and Data

To test each of these hypotheses, I rely on logistic regression. I include region-fixed effects to account for unobserved heterogeneity from region to region and I rely on robust standard errors clustered by country to obtain unbiased estimates from the panel data. To account for temporal autocorrelation, I follow [Carter and Signorino \(2010\)](#) and include the count of years since the last occurrence of violence along with its squared and cubed terms.

The dependent variable for one-sided violence (1975-2013) is derived from the UCDP one-sided violence dataset (Eck and Hultman 2007), which I convert from a count measure to a binary one, where 1 indicates at least one violent event in that given year and 0 none. The primary independent variable is the logged number of coethnic refugees a state hosts derived from the ER dataset (Rüegger and Bohnet 2015). As a check of robustness, I also include a binary measure of this variable where 1 indicates the presence of coethnic refugees and 0 indicates their absence.¹ In addition to coethnic refugees, I include a measure of the logged number of none coethnic refugees in each country to determine whether refugees with no ethnic ties to the country have a comparable effect on violence.²

I include the same set of controls for both models. I expect that rough and mountainous terrain hinders efforts to target civilians, particularly for government led forces. To control for this, I include a logged measure of the share of mountainous terrain. Furthermore, past studies have shown that various forms of internal conflict have become more prevalent and more deadly since the Cold War (Kalyvas 2001), so I include a dummy for the end of the Cold War. It is reasonable to suspect that one-sided targeting of civilians is more likely in the midst of civil war, so I include a dummy indicator for on going civil wars drawn from the ACD dataset (Wimmer, Cederman and Min 2009). I also include a number of social, economic, and political indicators, including the log of GDP/capita, the log of the infant mortality rate, and a dummy indicating whether a country experienced less than 2% growth in the previous fiscal year all drawn from the World Bank WDI dataset. Moreover, I control for whether a country is a democracy using the Polity IV coding of countries above the threshold of 6 (Marshall and Gurr 2014). Similarly, I control for regime stability by including the log of the number of years since a change of more than 2 points in a country's Polity IV score.

I also suspect that one-sided violence is least likely to occur in highly ethnically fractionalized societies and highly homogenized ones.³ In the former, if ethnic groups are too numerous, mobilization usually takes place along alternative lines (Esteban and Ray 2008), which may reduce the ability of warring parties to identify loyal and disloyal pools of civilians (Caselli and Coleman 2013), including refugees (see next chapter). Likewise, the salience of ethnicity is obviously low in homogeneous societies and for similar reasons I suspect homogeneous societies to also see less one-sided violence. To control for this curvilinear relationship where very low and very high levels of fractionalization are more peaceful than moderately fractionalized societies, I include both the commonly used measure of ethnic fractionalization derived from the Herfindahl

¹The results are robust to both measures so I present the results of the count variable only.

²I also restrict this measure to only include refugees from neighboring states. The results remain the same, thus, I report findings only for all refugees, irrespective of origin.

³I rely on a measure of "ethnic fractionalization" as opposed to "ethnic polarization" because past research has shown that polarization affects the level of conflict intensity. In contrast, fractionalization has been shown to increase the chances of conflict but not its severity. For more on this topic see Esteban and Ray (2008).

index and its squared term. Finally, all the variables in this analysis are lagged by one year. [Table 3.3](#) shows the descriptive statistics for each of these variables.

TABLE 3.3: Descriptive Statistics

Statistic	N	Mean	St. Dev.	Min	Max
One-Sided Violence _{Rebel}	5,185	0.105	0.306	0	1
One-Sided Violence _{Gov.}	4,994	0.059	0.235	0	1
Civil War	5,065	0.234	0.424	0	1
Post Cold War	4,963	0.590	0.492	0	1
Democracy	4,994	0.420	0.494	0	1
Slow Growth	4,600	0.307	0.461	0	1
Peace Years	5,026	12.541	10.003	0	35
Regime Durability	4,963	2.498	1.331	0.000	5.303
log(Refugees)	5,185	7.700	4.236	0.000	13.453
log(Ethnic Refugees)	4,745	2.246	4.390	0.000	14.950
Ethnic Fractionalization	5,095	0.416	0.292	0.001	0.925
log(GDP/capita)	5,065	1.303	1.345	-1.827	8.103
log(Population)	5,065	9.404	1.469	6.096	14.094
log(Infant Mortality Rate)	4,958	-0.135	1.039	-2.714	1.809
log(Mountainous Terrain)	5,095	2.149	1.393	0.000	4.421

3.4.3 Results

I test two separate models -one for one-sided violence committed by government forces and the other for one-sided violence committed by rebels, the results are shown in [Table 3.4](#).⁴ What I uncover is that the presence of refugees with ethnic ties to the host population significantly increases the chances of civilian targeting by rebel forces but not by government forces.

The first model fails to accept [H₃](#), which anticipates increased chances of civilian targeting by government forces as a response to the presence of coethnic refugees. This may suggest that government forces do not view refugees supporting their transborder kin as threats or, if they do, they may face other constraints in employing violence, such as international obligations to safeguard refugees. More likely, however, a selection effect in the destination choices of refugees accounts for this insignificant finding.

A civilian fleeing conflict faces three choices when caught in the midst of crossfire; 1) she must choose whether to stay with her family or to relocate them to a safer place 2) if she stays, she may take up arms or continue bearing the brunt of the war as a noncombatant but if she flees, she faces the additional choice to either migrate to safer regions of her own country as an

⁴Region-fixed effects are removed from the government led model because of convergence issues caused by perfect separation. This was not an issue for the other model, which retains the region-fixed effects.

TABLE 3.4: Ethnic Refugees and One-sided Violence

	(Government)	(Rebel)
(Intercept)	−3.82** (0.88)	−4.72** (1.00)
Regime Durability	−0.05 (0.09)	0.07 (0.10)
log(Refugees)	−0.02 (0.03)	0.07 (0.04)
log(Ethnic Refugees)	−0.01 (0.02)	0.05* (0.02)
Ethnic Fractionalization ²	2.67 (1.64)	4.50* (1.83)
Ethnic Fractionalization	−2.43 (1.38)	−3.55* (1.66)
log(GDP/capita)	−0.23 (0.16)	−0.44* (0.21)
Democracy	−0.60* (0.24)	0.29 (0.27)
log(Population)	0.24** (0.09)	0.26* (0.11)
log(Infant Mortality)	0.23 (0.28)	0.01 (0.28)
Slow Growth	0.66** (0.20)	0.29 (0.23)
log(Mountainous Terrain)	−0.02 (0.08)	−0.05 (0.08)
Civil War	1.31** (0.26)	2.20** (0.27)
Post Cold War	0.85** (0.32)	0.71 (0.37)
Peace years	−0.73** (0.11)	−0.97** (0.09)
Peace years ²	0.05** (0.01)	0.07** (0.01)
Peace years ³	−0.00*** (0.00)	−0.00*** (0.00)
Num. obs.	4398	4275
AIC	1193.8	1204
χ^2	1193.8	1207.6
Fixed-Effects	None	Region

Note:

*p<0.05; **p<0.01

internally displaced person (IDP) or to seek asylum across international borders as a refugee 3) if she makes the choice to lead her family to safety across international borders, she must finally decide where the safest place to flee is among her available choices.

That choice is partly constrained by geography; mountains, deserts, and bodies of water are harder and more dangerous to traverse. But that choice is also shaped by social and political dynamics. For example, there is evidence suggesting refugees, *ceteris paribus*, are more likely to flee to more democratic and less volatile neighbors than to autocratic and unstable ones (Schmeidl 2001; Moore and Shellman 2004, 2007; Rüegger and Bohnet 2015) recently provided empirical confirmation of a long standing suspicion that refugees tend to flee to countries where their transborder kin reside. Moreover, leaders of neighboring states may even limit entry of asylum seekers due to concerns over their own domestic stability (Neumayer 2005). As a result of a confluence of these various factors, refugees tend to migrate to countries whose governments are most supportive of their presence. This may include countries where the transborder kin of these refugees reside, such as Albania which hosts ethnic Albanian refugees from Kosovo. It may also include countries who share political fortunes with certain groups among the refugees, such as Qaddafi who actively opened his borders to refugees fleeing the Chadian civil war. If refugees tend to flee to countries ruled by governments that are more tolerant of and receptive to their presence then the association between the presence of transborder refugees and rebellion would be understated.

If refugees do migrate to countries where they enjoy shared ethnic and political ties with those in power, however, then opposition groups may view these incoming refugees as liabilities to be dealt with. If the incoming refugees not only share political ties with groups in power, but also ethnic ones, then rebels can more easily identify which refugees are potential recruits for their war effort and which are potential recruits for the war effort of their rivals. The second model tests this by regressing the use of one-sided violence by rebels on the logged number of coethnic refugees along with a host of control variables. These findings confirm H₂; rebels respond to the increased presence of transborder refugees by targeting noncombatants. There is substantial anecdotal evidence to back up this claim as well. For example, the news media has highlighted the persecution Syrian refugees face in Lebanon at the hands of ISIS sympathizers,⁵ how Boko Haram specifically targets IDP camps in Nigeria with suicide bombings and organized attacks,⁶ and Congolese rebels who target camps in Nord-Kivu among⁷ among others, which suggests refugees, like local populations, contribute to shaping the contours of conflict.

⁵See for example, <http://english.al-akhbar.com/node/21557>

⁶See for example, <http://www.theguardian.com/world/2016/feb/10/nigerian-refugee-camp-hit-by-double-suicide-bombing-boko-haram>

⁷See for example, <http://www.pbs.org/newshour/updates/africa-july-dec08-congo-1031/>

However, not all refugees are targeted in equal fashion. In fact, the [model](#) of rebel led one-sided targeting suggests none coethnic refugees have no statistically significant impact on the likelihood of one-sided violence. I have argued that warring parties have an incentive to "control the population" by undercutting the support base of their rivals. Doing so involves targeting civilians sympathetic to the cause of their rivals, an aim that is easier to achieve when society's divisions fall along ethnic lines. Refugees form an important support base for rival parties. Rebels know this so they target refugees, like civilians, as a means of undermining their rivals' capacity. Therefore, it is entirely expected that none coethnic refugees complicate efforts to distinguish friend from foe. That is not to say that the presence of none coethnic refugees is not destabilizing in other ways. Indeed, the literature I discussed earlier has typically focused on the relationship between conflict and refugees irrespective of ethnicity. For example, refugees, it has been argued, may contribute to economic pressures that increase the likelihood of conflict in the host state irrespective of their ethnic ties,⁸ but in the context of one-sided violence, it appears they exert no statistically significant impact.

Of the control variables in the [model](#) of rebel led violence, the log of population, ongoing civil war, and the post Cold War era all increase the probability of rebel-led one-sided violence. Moreover, the coefficient of ethnic fractionalization is negative and significant, while its squared term is positive and significant, which suggests only moderately fractionalized states are at risk of one-sided conflict (this is true of both government led and rebel led one-sided violence). Finally, the positive and significant coefficients for North Africa/Middle East and for sub-Saharan Africa identify these two regions as the most dangerous for civilians relative to the referent group of Western Europe, the US, and Canada.

It must be noted that the empirical observation that greater levels of coethnic refugees are associated with greater levels of civilian targeted violence does not by itself establish that refugees are necessarily the targets of such violence. It is perfectly reasonable, for example, to suggest that the association between coethnic refugees obtains because they are either actively involved in the violence or because they embolden their transborder kin to rebel. However, if we are to believe that refugees do tend to flee to countries that are more inviting of their presence, and we have every reason to believe they do, then the association between coethnic refugees and government led, not rebel led, violence should be significant, which it is not. In other words, if refugees flee to countries with more accepting governments and if there existed a refugee warrior phenomena that was fueling one-sided violence, then we would observe government allied forces targeting more civilians, as opposed to rebel forces targeting them. In the concluding section, I highlight a number of observable implications of these findings and discuss approaches to testing them, which I carry out in the next [Hchapter](#).

⁸See for example, the Regional Center for Strategic Study's work on this topic at <http://www.rcssmideast.org/en/Article/144/Syrias-refugees-burden-neighboring-countries-economies-#.VxdCvdCNvt8>

3.5 Conclusion

In the previous [chapter](#), I showed that the presence of refugees increases incidents of conflict in Lebanon, particularly in districts with volatile ethnic relations. I accounted for this finding with the suggestion that refugees alter the ethnic balance of power between rival ethnic groups, which increases opportunities and incentives to engage in conflict. In this chapter, I further explore this proposition using a global dataset of all ethnic groups. The results failed to identify a link between coethnic refugees and the onset of ethnic armed rebellion, which suggests one of three things; either the association between the two was understated because refugees flee to countries where there transborder kin are in government and not opposition, coethnic refugees contribute to conflict severity but not to the onset of conflict itself, or coethnic refugees contribute to other forms of conflict all together, such as one-sided violence.

I tested this latter possibility using a global country-year dataset of one-sided violence. The results of this latter analysis point to a link between the presence of coethnic refugees and rebel led one-sided violence. That is, coethnic refugees are associated with violence not because they are perpetuating it, but because they are victims of rebels groups who target them. This is a major finding that frames the link between refugees and violence in ways yet to be explored by the field. For example, of the six major works in the field that quantitatively assesses the relationship between refugees and conflict, not a single one attempts to theoretically or empirically link increases in violence to attacks on civilians, including refugees themselves, despite the abundant anecdotal evidence reported in mainstream and alternatives media sources.⁹ In fact, theories tying refugees to conflict in each of these articles focus on how refugees produce dividends in the form of recruits for armed rebels. Much of the work in this dissertation relies on this same theoretical understanding but argues that this is precisely what makes refugees valuable targets for repression. Thus, the findings of this chapter provide a framework for thinking about refugees and their role in conflict that departs from conventional views of refugees as "warriors" in conflict ([Zolberg, Suhrke and Aguayo 1992](#); [Adelman 1998](#)).

If refugees are a source of power in domestic politics that can advantage one group over its rivals, then rebels have an incentive to control the population of their country as a means of bolstering their own support base and undermining the support of their rivals. This logic of population control entails targeting civilians and refugees who support their rivals' claims to power. Doing so, however, can be costly if rebels cannot distinguish between noncombatants who sympathize with them and those sympathetic to their rivals. But if society is mobilized along ethnic lines, then discriminating between friend and foe is easier. In the context of refugee

⁹See for example, [Salehyan and Gleditsch \(2006\)](#); [Salehyan \(2008\)](#); [Forsberg \(2009\)](#); [Bohnet \(2012\)](#); [Fisk \(2014\)](#); [Shaver and Zhou \(2015\)](#). In fact, [Zolberg, Suhrke and Aguayo \(1992\)](#) coined the term "refugee warrior" decades ago to describe forced migrants who take up arms in their destinations of refuge.

flows, this suggests that ethnic refugee should increase the chances of one-sided violence but non-ethnic refugees should not. This much, at least, has been confirmed by the empirical analysis of the current chapter but there are other observable implications of this theory as well. Some of these observable implications have been studied by other scholars. For example, if rebels have an incentive to augment their public support while undercutting that of their rivals and if doing so is easier when society is divided along ethnic lines, then we should expect civilian targeting to be more severe during ethnic conflict than nonethnic conflict.

Indeed, there is much work to back up this claim ([Valentino, Huth and Balch-Lindsay 2004](#); [Eck 2009](#); [Caselli and Coleman 2013](#)) but there are other observable implications that have yet to be examined. In the following [chapter](#) I look one of these. I suggest that if rebels have an incentive to target noncombatants and if doing so is easier under conditions of ethnic conflict, then we should observe a statistically significant difference between the patterns of forced migration resulting from ethnic civil wars and those resulting from nonethnic conflict. I hypothesize that ethnic civil wars produce a greater share of refugees relative to internally displaced persons (IDPs) than nonethnic civil wars and I find empirical evidence in support of this claim.

The theory of the logic of population control can account for this finding. Civilians fleeing violence have the choice to flee to safer regions of their own countries or to seek asylum in foreign states. If all the civilians fear is the crossfire of conflict, they will be more inclined, I argue, to stay as close to their homes as possible without placing themselves and their families at risk. Typically, this means relocating to other regions of the same country. However, if the combatants have an incentive to specifically target civilians, then civilians will find it in their interest to relocate to places warring parties cannot follow, which in the context of an actually functioning Westphalian system of states means seeking asylum across international borders as registered refugees. The next [chapter](#) will explore these propositions in more detail.

Chapter 4

The Logic of Population Control

Abstract

Why do some civil war torn countries produce more refugees relative to their internally displaced population and others displace more of their population internally than across borders? Surprisingly, the relationship between internally displaced persons and conflict has been woefully underexplored. The aim of this chapter is to fill this gap in the literature. Using a panel dataset of civil conflicts by country-year from 1993-2010 and a two-step Heckman selection model, I show that civil wars fought along ethnic lines produce greater refugee flows relative to IDP flows than non-ethnic civil wars. I account for this finding by relying on insights drawn from the previous chapters. Specifically, I argue that in conflicts where combatants are recruited along ethnic lines, ethnic markers allow for less costly and more discriminate targeting of rival civilian populations, which in turn increases the share of forced migrants who seek refugee across borders relative to those displaced internally.

4.1 Introduction

Much of the literature on refugees and conflict contagion has focused on the conditions under which refugees are more or less likely to spread violence across borders. There is some evidence, for example, that the effects of refugees on the spread of violence is conditioned by a number of factors, such as shared ethnic ties with the host population, resource scarcity, and political instability in the host state. Specifically, this dissertation has argued that refugees contribute to violence in two primary ways. First, refugees alter the existing balance of power between competing ethnoreligious groups at the substate level, which can result in increased levels of violence in the host state if such relations are polarized to begin with. Second, coethnic refugees

also increase the likelihood of one-sided violence, especially at the hands of non-state actors, suggesting that their presence is a potential threat to certain rebel groups. The logic in both these circumstances is uniform; refugees represent both a threat and a boon to rebel groups. As much as coethnic refugees are an attractive target for recruitment and resource extraction for one rebel group, their presence is viewed as a threat by competing rebel groups.

What is less clear in the literature, however, are the dynamics that contribute to the emergence of refugees in the first place. In many ways, we've put the proverbial cart before the horse. We know how conflict refugees contribute to the spread of violence yet we do not know which conflicts are most likely to result in a large exodus of refugees in the first place. From the perspective of the conflict contagion literature, we know which countries are at greatest risk of contagion but we do not know which countries are the most *contagious*.

While the forced movement of people is a feature endemic to most conflict-ridden states,¹ the proportion of civil war-torn states with active refugee *flows* represent only 51% of all conflict years between 1993-2011.²

The range in the number of forced migrants fleeing conflicts is also non-trivial. For example, the 1994 civil war and genocide in Rwanda resulted in the flight of a little over 1.8 million refugees. Similarly, in 2006, the intercommunal violence and civil war that plagued war-torn Iraq resulted in nearly 1.2 million people fleeing their country. Compare those annual figures, with the number of refugees that fled Tajikistan's civil war in the 1990s –only 5,000 refugees were displaced over the course of five years. Of course not all victims of forced movement in conflict become refugees –a good portion of them are displaced internally. In fact, in many cases the forced displacement of victims of war within borders exceeds the displacement of victims of war across borders. Take for example Pakistan's civil war against the Taliban insurgency, which resulted in the internal displacement of over 1.7 million people in the year 2009 alone. This stands in great contrast to the 2,279 individuals who sought and found refuge *outside* Pakistan's borders that same year.

Why do some civil war torn countries produce more refugees relative to their internally displaced population and others displace more of their population internally than across borders? Surprisingly, the relationship between internally displaced persons and conflict has been woefully underexplored. The aim of this chapter is to fill this gap in the literature. Using a panel dataset of civil conflicts by country-year from 1993-2010 and a *two*-step Heckman selection model, I show that civil wars fought along ethnic lines produce greater refugee flows relative to IDP flows than non-ethnic civil wars. I account for this finding by relying on insights drawn from

¹For example, between 1993 and 2011, only one country enduring conflict featured no refugees living outside its borders, Papua New Guinea between 1993-1994.

²Active refugee flows refer to the number of refugees forced to flee in a given year from a given country, as opposed to refugee stock, which is a count of the cumulative total of refugees.

the previous chapters. Specifically, I argue that in conflicts where combatants are recruited along ethnic lines, ethnic markers allow for less costly and more discriminate targeting of rival civilian populations, which in turn increases the share of forced migrants who seek refugee across borders relative to those displaced internally.

In what remains of this chapter, I introduce a theory, which I label the “the logic of population control”, that accounts for variation in levels and patterns of forced migration resulting from civil wars. This logic of population control is grounded in theoretical contributions from three disparate research programs –ethnic conflict, civilian targeting, and forced migration. In the sections that follow, I briefly review the literature in each of these research agendas in that order, I provide a stylized narrative of the theory, and I present my methodological approach to testing the observable implications of the theory. Finally, in the last two sections I provide an interpretation of the results and conclude with a discussion that ties these findings to the conflict contagion research agenda.

4.2 Literature Review

4.2.1 Ethnicity and Ethnic Civil Wars

Although the literature disaggregates civil wars in a number of different ways, such as “old vs new wars” (Kalyvas 2001), “territorial vs governmental” (Gleditsch et al. 2002), “symmetric, non-symmetric, conventional, and irregular” (Kalyvas 2005), etc., perhaps the most widely employed typology of civil war contrasts ethnic with non-ethnic violence. This study embraces this latter approach. In general, justifications for analytically separating ethnic and nonethnic conflict focus on motivations, structural conditions, and opportunities that differentiate one form from the other.³

Justifications based on motivation typically highlight the unique set of ethnic grievances that result from differential treatment of groups. For example, Sambanis (2001) finds evidence that ethnic war is waged in response to political grievances, while nonethnic war is associated with lack of economic opportunity. In other words, ethnic discrimination is a uniquely motivating factor in rebellion. Others argue that ethnic and religious discrimination produce grievances that result in conflicts that are more difficult to manage. For instance, Hassner (2009) suggests that religious grievances that result in territorial conflict are more intractable and difficult, if not impossible, to resolve due to the motivating beliefs of those involved over the sacredness of the territory being fought over. Similarly, Fearon (1995, 1998, 2004) argues that ethnic wars

³This distinction has not been accepted wholesale in the field. See for example, ? who argues ethnic conflicts resemble other forms of violence in so far as they are waged by small groups of combatants purported to “fight and kill in the name of some larger entity.”

produce commitment problems that result in extended duration of warfare relative to nonethnic conflicts. Furthermore, [Kaufmann \(1996\)](#) contrasts “ethnic civil wars” with “ideological civil wars”. The key difference between these two forms of war he suggests is the level of flexibility of individual loyalties. Where loyalty in ideological conflicts is quite fluid, in ethnic conflicts it is far more rigid. According to Kaufmann, the rigidity of loyalty is what makes ethnic conflicts particularly difficult to resolve peacefully.

In contrast to motivation-based explanations, justifications based on the different structural conditions of ethnic and nonethnic warfare rely on the ethnic composition of conflict prone societies. Researchers who proffer these justifications identify variation in levels of ethnic polarization and/or fractionalization as factors that condition warfare in unique ways that require tailored approaches ([Reynal-Querol 2002](#); [Garcia-Montalvo and Reynal-Querol 2004](#); [Bhavnani and Miodownik 2009](#)).

Finally, opportunity justifications typically focus on the conditions unique to mobilization when ethnicity is salient and highly politicized. These sorts of explanations look to the ways that ethnic markers shape the ability of groups to mobilize for warfare. ? argue ethnic markers help enforce group membership by reducing “free-riding”. In homogenous societies members of the losing group can easily pass themselves off as members of the winning coalition but in ethnically heterogeneous societies ethnic markers make this task far more difficult. The fact that leaders in homogenous societies understand this *ex post* dilemma reduces their incentives to mobilize *ex ante* in none-ethnically salient states. These ethnic markers not only reduce the free-riding problem, they also allow groups to easily identify and target loyal populations for their recruitment efforts. Thus, ethnicity plays a role in reducing the coordination costs associated with mobilization. These reduced barriers to mobilization, of course, result in greater risks of conflict escalation ?. While ethnic markers provide groups with more efficient means to mobilize, they also provide combatants with more effective means of targeting pools of populations loyal to rivals. In effect, the opportunity structures of ethnic civil wars increase incentives to target civilians ([Valentino, Huth and Balch-Lindsay 2004](#)).

The distinction between ethnic and nonethnic conflict made in this study is motivated by this very phenomenon. Thus, I embrace the opportunity-based justification for analytically separating ethnic from nonethnic conflict. If ethnic markers provide combatants with better opportunities to target the loyal population base of rival groups, then it stands to reason that civilians are at a greater risk of victimization in ethnic civil wars. The resulting atmosphere of violence should increase the incentives of forced migrants to seek refugee across international borders (as opposed to hiding among the general population). In the following section, I discuss the extant literature on the targeting of civilians during combat and tie it to the literature on forced migration and ethnic conflict.

4.2.2 Civilian Victimization

Civil war literature has identified a number of factors that contribute to the victimization of civilians during conflict, including autocratic regime types (Engelhardt 1992; Harff 2003; Valentino, Huth and Balch-Lindsay 2004), use of guerrilla tactics by rebels (Valentino, Huth and Balch-Lindsay 2004), and desperation to win (Downes 2006). Scholars studying this topic have also embraced the ethnic-nonethnic distinction in examining civil wars and their analyses suggest ethnic conflict increases the chances of civilian targeting relative to nonethnic conflict (Downes 2006; Fjelde and Hultman 2010). Valentino, Huth and Balch-Lindsay (2004) argue that because “it is more difficult for individuals to disguise their ethnicity than their political affiliation” combatants are better able to discriminate between friendly and hostile civilian populations. In the absence of ethnic cleavages, combatants find it difficult to distinguish friend from foe; as a consequence, nonethnic civil wars actually result in less civilian targeting than ethnic ones (*ibid*). This has important implications for the patterns of forced migration that result from domestic warfare. The following section identifies these observable implications after briefly reviewing recent quantitative literature on forced migration.

4.2.3 Forced Migration

A number of significant factors have been shown to increase the risk of forced migration in a given a country. Typically, these fall into one of two categories –push or pull factors. Push factors refer to characteristics and conditions of countries that force populations to flee their homes to safer destinations abroad. Examples of push factors include natural disasters (Drabo and Mbaye 2011) and various forms of violence (Schmeidl 2001; Moore and Shellman 2004; Davenport, Moore and Poe 2003). Pull factors refer to attractive neighborhood characteristics that make movement across borders less costly than seeking refuge within borders (or staying put all together). These include the regime type of destination countries (Moore and Shellman 2007), shared ethnic affiliation in countries of destination (Rüegger and Bohnet 2015), and hospitable neighboring geography (Moore and Shellman 2007). An analysis of the factors that contribute to one form of forced migration (international refugees) or another form (internally displaced persons) must incorporate both of these elements. Surprisingly, the relationship between external and internal displacement has only received scant attention in the field thus far.

The only work to date that I am aware of that quantitatively compares refugee flight and IDP movement to one another is Moore and Shellman (2007). They employ a two-step Heckman model on a global panel analysis of country-years between 1976-1995. Their analysis suggests that levels of violence in neighboring states increases the proportion of IDPs flows relative to refugee flows. In their formulae, victims of displacement weigh the dangers they perceive at

home against those they see in their potential points of destination. When conditions in their potential points of destination are more favorable than the conditions they face at home, they are more likely to migrate (and *vice versa*). The current analysis embraces both the methodology and theoretical foundations of Moore and Shellman’s article. However, it departs from their approach in two ways. First, where Moore and Shellman’s narrative emphasizes the agency migrants enjoy in determining their own choice of destination, the present analysis examines the role that combatants play in limiting that agency and influencing that choice. Second, where Moore and Shellman examine “characteristics of countries” that affect patterns of forced migration, the present analysis also examines the *characteristics of conflicts* that affect that phenomenon. This latter point is not trivial and, in fact, represents a major departure from the methodological approach of Moore and Shellman and much of the existing literature on forced migration.

Instead of examining patterns of forced migration for all countries as others have done, I am specifically interested in the patterns of forced migration under the strategic environment that victims and combatants of civil war find themselves. Therefore, I restrict my analysis to country-years experiencing at least one civil war. I argue this exclusion criteria is justified because the choice to stay, flee to other regions, or to seek refugee in other countries is inherently different for those in an environment of organized warfare than those fleeing economic hardships, natural disasters, or other forms of political violence.⁴ That difference, I contend, arises from the unique incentives combatants (i.e. rebels and governments) have to control the flow of population within and between their territories and regions controlled by their rivals. Furthermore, I suggest the type of war being waged shapes these very incentives, perhaps to a great deal. The next section explains why. In what follows, I introduce the “logic of population control”, I establish the assumptions of the theory, I introduce the relevant actors and their interests, I formalize the opportunity structures that govern their behaviors, and I provide a stylized account of the theory at work.

4.3 The Logic of Population Control

The logic of population control is based on the premise that combatants have an incentive to control populations loyal to them and to undermine the control their rivals enjoy over their own loyal populations. In the previous chapters, I examined the role of ethnicity in fomenting conflict in regions of Lebanon heavily populated by Syrian refugees. Specifically, I argued that

⁴This exclusion criterion can potentially introduce bias into the model’s estimates if civil war-torn and peaceful states differ in unobservable ways either related to the likelihood of forced migration or to the proportions of forced migrants that are refugees (or IDPs). Please see the Data and Methods section for a more in-depth discussion of the selection problem and the tools I use to manage the issue.

refugees alter the balance of power between rival ethnic groups, which can result in increased mobilization efforts towards organized violence. My analysis showed that 1) higher refugee numbers increase conflict and 2) ethnically polarized regions are particularly susceptible to this threat. Refugees, like the local population, are a vital resource for the groups that they share ethnic affiliations with. Therefore, rival groups view refugees of rival ethnic groups as a threat. In Lebanon, this manifested itself in one-sided attacks on refugees, organized attacks between rival groups, and increased intercommunal violence amongst the local population.

The analyses from the previous chapters suggest that control over populations is an important goal for combatants. Asserting control over loyal populations and undercutting the support their rivals enjoy provides combatants with a comparative advantage in mobilization efforts and in resource extraction. In the same way that groups have an interest in targeting refugees of rival ethnic groups, combatants also have an incentive to target the local populations of rival ethnic groups. Thus, for the very same reason that the presence of *coethnic* refugees is more destabilizing in a host country than the presence of *non-coethnic* refugees, civil wars fought along ethnoreligious divisions are also more destabilizing than those fought along other societal divisions (i.e. class, ideology, etc.). They are destabilizing in ways that are particularly destructive towards civilians. In the previous section, I discussed the literature on civilian targeting that identified ethnic civil wars as the most violent form of civil war that civilians endure. The theory introduced here examines how this logic of population control influences the patterns of forced migration that result from ethnic and nonethnic conflict alike. Next, I turn my attention to the assumptions of the theory.

4.3.1 Assumptions

I begin with a number of important assumptions. First, although civilians make a ‘decision’ to flee (either within or across borders), ultimately “forced migration” invariably entails a level of coercion that really places more agency in the hands of perpetrators of the violence (combatants) than in the hands of its victims (refugees and IDPs). Therefore, I assume that combatants maintain some degree of influence over the decision of civilians to stay, to flee to other regions of the country, or to seek refugee in other states.

My second assumption is that population control is an important aim of combatants in civil war. Control over populations allows parties to a conflict to extract human and material resources in the form of soldiers, field doctors/nurses, taxes, and general economic production. Population control also allows combatants to homogenize their populations in support of their war aims by expelling, detaining, or killing sympathizers of the opposing group. For the very same reason that a controlled but robust and supportive population is critical for success in conflict, warring

parties have an interest in undercutting the population support their rivals enjoy. One common way of doing so in the midst of warfare is directly targeting civilians.

My next assumption rests on the difference between war-torn states where ethnicity plays a salient role in the society's divisions and where it does not. The costs of and barriers to population control in ethnically salient conflicts are, as already discussed, less than in nonethnic conflicts because ethnic markers allow groups to better discriminate between supporters and those sympathetic to opposing groups. Therefore, I contend that different *types* of civil wars produce different logics of population control.

Furthermore, many civil wars increase the threat of bodily harm to civilians and they react to this threat by fleeing their homes. They can either flee to other regions of their own country or attempt to seek refugee across the border in neighboring states. I assume, all else equal, movement across international borders is more costly than internal displacement. Therefore, civilians should prefer to relocate as close to their original location of residency as possible without exposing themselves to high risks of victimization. Ultimately, the decision to relocate internally as opposed to seeking asylum elsewhere, wrests on whether the civilian thinks his/her probability of victimization is higher outside the country than inside. That calculation, I argue, is influenced by the likelihood that combatants will target civilians, which is itself influenced by the type of civil war waged.

My final assumption is that rightly or wrongly, civilians are viewed by rebels (or governments) representing rival groups as a potentially threatening resource at the disposal of their enemies. Under conditions of non-ethnic conflict the tools of targeted repression become blunted because governments (or rebel groups) find it difficult to distinguish between loyal and disloyal pools of civilians. Because mass repression can undercut support among previously loyal populations, rival victims of war are better able to seek refugee undetected within the borders of the state by hiding among populations that warring groups are hesitant to target or find it difficult to target effectively.

In contrast, the conditions of ethnic conflict produce a different logic of repression and flight. Ethnic markers provide warring groups with the ability to not only identify supportive populations but hostile ones as well, a point ? makes when he suggests that combatants "can treat all members of the other ethnic group as enemies without risk of losing a recruit" (21). In effect, conditions of ethnic conflict limit the domestic destination options available to fleeing victims of war.

4.3.2 Actors and Interests

The notion that combatants have incentives to control the type of forced migration that results from their participation in ongoing violence I label the theory of population control. The narrative of this theory focuses on three actors –rebels, governments, and civilians. Each of these groups has their own interests as well. Rebels want to limit the resources of the government and one way to do so is to target populations loyal to them. Similarly, governments want to undercut the support base of opposing rebel groups, so they too benefit from targeting civilian populations loyal to their rivals. At the same time, both governments and rebels have an incentive to safeguard loyal populations. Therefore, overzealous targeting of civilian groups that puts their own potential supporters at risk is costly. Finally, civilians too are self-interested actors whose primary goal is to reduce their personal likelihood of persecution at the hand of combatants.

Let T_l denote the probability of targeting loyal civilians and let T_d denote the probability of targeting disloyal civilians. Line 4.1 identifies the incentives versus the constraints combatants face under conditions of nonethnic civil war,

$$T_l = T_d \quad (4.1)$$

That is, when groups are mobilized along nonethnic lines, the probability of targeting loyal civilians should be roughly equal to the probability of targeting populations loyal to rivals because distinguishing between friend and foe is more difficult without observable markers that aid in differentiating one from the other. In contrast, under conditions of *ethnic* conflict, the likelihood of targeting disloyal civilians is greater than the likelihood of targeting one's own population base,

$$T_l > T_d \quad (4.2)$$

The balance between T_d and T_l is influenced by the ability of combatants to target rivals' civilian support discriminately. When groups are mobilized along ethnic lines, the dividends, from attacks on civilians is higher than the risks associated with civilian victimization because there are fewer costs to targeting civilians discriminately. However, when fighting takes place along the lines of non-ascriptive identities, such as class, fewer observable markers exists that reveal the loyalties of the civilian population. Under these conditions, combatants find the chances of targeting loyal civilians, T_l and disloyal ones, T_d , roughly the same. Let C denote the costs associated with targeting civilians and B the benefits. When $T_d > T_l$ then,

$$C > B \quad (4.3)$$

In other words, when the probability of targeting disloyal civilians (T_d) is higher than the probability of targeting loyal ones (T_l) then the costs (C) of targeting civilians is less than the benefits (B) accrued from the tactic. If, V_e , represents the level of violence against civilians during ethnic civil war and V_n the level of violence during nonethnic civil war, then,

$$V_e > V_n \quad (4.4)$$

Line 4.4 indicates that under conditions of ethnic conflict violence against civilians is greater than under conditions of nonethnic conflict. Civilians also make a cost-benefit analysis. P_o refers to the probability of perceived victimization in a civilian's country of origin and P_a refers to the probability of perceived victimization in a civilian's potential destination of asylum. When P_o are equal P_a ,

$$P_o = P_a \quad (4.5)$$

civilians will choose to relocate to domestic destinations. That is, when the perceived chances of being targeted is the same at home that it is abroad (or in the process of traveling abroad) then civilians will elect to stay within the borders of their country. Likewise, if the perceived probability of persecution abroad is higher than at home,

$$P_o < P_a \quad (4.6)$$

then civilians will seek refugee in regions of the country they find safer than the battle grounds from which they escaped. On the other hand, if the perceived probability of persecution at home is higher than the perceived likelihood of victimization abroad (and in the process of traveling abroad),

$$P_o > P_a \quad (4.7)$$

then civilians will seek refugee across international borders. Let R represents the number of refugees fleeing civil war and I represent the number of IDPs fleeing civil war. F_e and F_n represent the share of forced migrants that are refugees for ethnic and nonethnic civil wars respectively.

$$F_e = \frac{R}{R + I} \quad (4.8)$$

$$F_n = \frac{R}{R + I} \quad (4.9)$$

Holding all else equal, combatants find $C < B$, when $T_d > T_l$ is true. And when $T_d > T_l$ is true, civilians will calculate $P_o > P_a$. It is the contention of this chapter that $T_d > T_l$ is true during ethnic conflicts more so than during nonethnic conflicts. If $P_o > P_a$ then, all else equal,

$$F_e > F_n \quad (4.10)$$

In other words, during conditions of nonethnic conflict the share of forced migrant flows that are composed of refugees is fewer than during conditions of ethnic conflict. Thus compared to nonethnic civil wars, civil wars characterized by ethnic cleavages are more likely to produce conflict environments that push civilians to seek asylum among foreign populations than to find shelter among their own.

4.3.3 Stylized Narrative

The logic of population control, thus, identifies two patterns of flight. When the costs of targeting loyal populations are higher than the benefits accrued from employing this tactic, $C > B$, civilians become bystanders in danger of crossfire. As such, their pattern of flight can be characterized as one that avoids regions of high intensity warfare between rival groups. If $C < B$, however, civilians are not mere bystanders in war. In effect, they become targets of war and their flight patterns will reflect their intent to not only avoid regions of high intensity conflict but also peaceful places either controlled by rival adversaries or at risk of control (or targeting) by such groups.

From these two sets of equivalencies we are left with a stylized narrative that describes the conditions under which civil wars produce more or less refugees as a share of total forced migrants. Imagine for a moment a head of a household caught in the crossfire between rebel and government forces. She faces the choice to stay with her family, flee with them to more peaceful regions of the country, or make the potentially dangerous trek across international borders to safety. Therefore, she makes two choices. First, whether to stay or flee and, second, once on the move, whether to relocate to other regions or to seek refugees across the border. Her first decision is simple; she will relocate with her family in attempt to evade the threat of

crossfire. Her second choice, however, depends on the perceived likelihood that the violence will follow her and her family to their choice of destination.

The perception of this likelihood is influenced by the deliberate actions of the warring parties because at the very moment she is deciding whether to simply flee or seek actual asylum, group decision makers are faced with a choice whether to target her and her family as they flee. On the one hand, if ethnic markers reveal the direction of her loyalty then the choice to target is simpler to make. If the civilian's ethnic markers reveal her to be a of a rival ethnic group then the risks associated with targeting her are lower and decision makers will likely make the choice to target. If the civilian markers identify her as supportive, then rival groups may target her instead.

On the other hand, if ethnic markers do not coincide with the cleavages of warfare, then choosing to target such a civilian is costly –you may or may not have just targeted a supporter. Thus, under conditions of nonethnic conflict the propensity to target civilians is reduced and civilians like that head of the household use this information when deciding where to flee. If they experience deliberate targeting by combatants, they will take this as a cue that the violence will follow them to where they flee. Therefore, civilians will be more likely to flee to regions outside the reach of warring parties (i.e. outside the country). But if they are not targeted and, as such, view the threat they face in the conflict as incidental, then they will be more inclined to seek refugee in places they consider safe from crossfire. The following section posits a set of hypotheses derived from the discussion above.

4.4 Hypotheses

In the previous two sections I introduced the logic of population control. In this section I discuss a number of observable implications of this theory and formally present a set of hypotheses, which I empirically test in subsequent sections. The most basic observable implication of the theory is that in the midst of civil war, refugees respond to direct threats to their personal safety by relocating either to other regions within their own countries or to safe regions in other countries. From this I derive my first hypothesis:

H₁: *Controlling for all other factors, one-sided violence against civilians increases the probability of forced migration among civil war torn states.*

Moreover, for the two-step Heckman correction to work properly at least one variable must act as an “instrument” for the effect of forced migration; that is, it must predict forced migration without affecting the composition of forced migrants. In this case, one-sided violence acts as the instrument. If one-sided violence has a significant *independent marginal* effect on the

composition of forced migration then its role as an instrument is compromised because it affects the probability of forced migration to begin with. But if, as I suspect, one-sided violence increases the probability of forced migration (H_1) without affecting the composition of forced migration (i.e. the share of forced migrants that are refugees), then introducing an interaction of one-sided violence and civil war type (as is suggested by (H_3) and (H_4) below) should pose no problems for estimation even if the interaction itself is significant. Thus I hypothesize the following null effect:

H₂: *One-sided violence has no independent effect on the proportion of forced migrants that are refugees.*

The key proposition presented in the section discussing actors and their interests is that combatants have incentives to target civilians and when they do, refugees have a choice to flee to other regions of their home countries or to flee to safety across international borders. If the war is fought along ethnic lines, then combatants can more effectively target these civilians based on observable ethnic markets. Realizing this, civilians in the midst of ethnic conflicts will find it less attractive to seek refugee in their own countries, which brings me to my next hypothesis:

H₃ *Controlling for all other factors, compared to nonethnic civil wars, civil wars characterized by ethnic cleavages produce more refugees as a share of total forced migrants.*

As discussed earlier, ethnic markers increase the incentives combatants have to target civilians because it allows for more discriminate targeting and previous research has verified this claim. Therefore, ethnic markers act as a sort of intervening variable where the effect of one-sided violence on migrant patterns is heightened if the conflict is waged along ethnic lines. In fact, I suspect that one-sided violence against civilians increases the share of refugees relative to IDPs but only under conditions of ethnic conflict. Thus, the effect of civil war type on composition of forced migrants is conditional on the presence or absence of one-sided violence:

H₄: *Controlling for all other factors, the positive effect of ethnic civil wars on the share of forced migrants that are refugees is conditional on the presence of one-sided violence.*

If the civilians being targeted find themselves in nonethnic conflict, however, then I expect they will be more inclined to relocate to regions within their own country than to neighboring states primarily because of the difficulty in traversing international borders. If it is easier to hide among the local population, which it is under conditions of nonethnic conflict, civilians will do so. From this I derive my next hypothesis.

H₅: *Controlling for all other factors, one-sided violence against civilians under conditions of nonethnic conflict should decrease the share of forced migrants that are refugees.*

H₆: *Controlling for all other factors, civil war type should not affect the share of forced migrants that are refugees absent one-sided violence.*

Similarly, I do not expect ethnic conflict to increase the share of forced migrants that are refugees when one-sided violence is not a feature of warfare. In fact, if the only mechanism by which ethnic conflict affects patterns of forced migration is the level of one-sided violence it produces, then it stands to reason that ethnic civil wars free of one-sided violence should have no significant effect. (H₆) reflects this expectation that the effect ethnicity on the share of migrants that are refugees is conditional on the presence of one-sided violence against civilians. The next section introduces the methodological approach, estimation techniques, and data.

4.5 Data and Methods

4.5.1 Methodological Approach

The empirical strategy adopted in this chapter is primarily driven by the methodological challenge posed by selection bias in at least one stage of the analysis. The problem of sample selection is a form of omitted variable bias that arises from a nonrandom selection of data. When a subset of the data is systematically excluded due to a particular factor then exclusion of the subset can bias estimates.

This problem can emerge as an artifact of the research design or when subjects self-select into certain groups. As an example of the latter, if a researcher is interested in the effect of drug use on mental illness, simply regressing mental illness on past drug use will yield bias results if individuals use drugs to self medicate. The researcher may identify more drug use with greater mental health problems but the relationship may very well be overstated if we think the reason for that higher drug use had to do with greater levels of mental illness to begin with. This is also known as endogeneity bias because the selection into treatment groups is endogenous to the outcome. In regards to the former, if a researcher is interested in examining the relationship between education and wage offers but only has access to wage offers of individuals currently employed then the factors that predict participation in the labor force may bias the relationship between levels of education and wage offers, specially if labor force participation and wages are related.

Both these forms of bias may exist in the current study and their threats to inference are to be taken seriously. I begin with a discussion of subject-self selection (also known as endogeneity or treatment selection) referred to above in the anecdote of the study of drug use and mental illness. I have already established that ethnic conflicts better facilitate combatant recruitment and mobilization. Not only do ethnic markers allow leaders to better commit to the rank and

file, they provide leaders with assurances that their spoils of war will not be diluted by free riders *ex post*. Furthermore, these ethnic markers allow combatants to more effectively target civilians loyal to rival groups. Therefore, there exists at least some nominal benefit to mobilizing along ethnic lines. However, leaders may face countervailing factors that push them to mobilize along alternative cleavages instead. If leaders are capable of determining the societal fissures along which war is waged and for some unobservable reason(s) leaders that choose to mobilize their rank and file along nonethnic lines also happen to be leaders that are less willing to target civilians during warfare, then the effect of ethnic civil wars (in contrast to nonethnic wars) on the proportion of forced migrants that are refugees will be overstated.

One way to address this threat to inference is to first model the likelihood that civil war is fought along ethnic lines and then move on to examining the outcome of interest. This is may be unnecessary, however, because 1) I contend that while leaders certainly have an interest in mobilizing their rank and file according to the ways they see most beneficial to their cause, their actual ability to do so in meaningful ways is very limited. In fact, the very politicization of ethnicity is a process that takes shape over many years and under the influence of many factors and 2) the research design faces other antecedent biases in the chain of selection that are more important to address. I turn to those next.

Although, I begin with a total sample of all country years between 1993-2010 (based on data availability), I am only interested in the pressures that forced migrants face in the midst of civil war. Thus, I subset the data to country-years experiencing an ongoing civil war. An argument could be made that excluding all none civil war states will introduce selection bias in the final estimates. But this misses the point; exclusion of observations based on a particular selection of a population of interest does not induce bias on its own. In fact, [Wooldridge \(2010\)](#) asserts “sample selection can only be an issue once the population of interest has been carefully specified” (551). He suggests that if the researcher is interested in a subset of a larger population then the appropriate approach is to specify a model for the part of the population based on randomly selected data from that subset. In this study, I am only interested in examining the strategic environment victims and combatants of civil wars find themselves. Therefore, selection based on my “population of interest” –civil wars –should not influence my results.⁵

⁵Indeed, the canonical heckit correction method ([Heckman 1979](#)) for identifying and addressing selection effects fails to uncover a selection bias produced by exclusion of non-civil war states; the inverse Mill’s ratio of the bivariate probit estimation does not attain significance. Had selection bias been revealed at that stage as it has been in a subsequent stage, then a bivariate probit model would be used to identify both selection effects (selection into civil war states and, given a set of civil war-torn states, selection into those that produce forced migration) and the inverse Mill’s ratio obtained from both selection equations would be included as additional regressors in the outcome equation that estimates the effect of civil war type on pattern of forced migration. Such an approach, a multi-stage selection model, would be warranted if the inverse Mill’s ratios (identified below) of both selection equations attain significance. However, this method identifies bias only in the second stage of selection, selection into civil war-torn states that produce forced migration. In other words, excluding civil war free country-years from the sample does not bias the results, while exclusion of states that do not produce forced

Next, I need to examine whether some civil war torn states produce more refugees relative to IDPs than others and accounts for that difference. However, not all civil war-torn states experience forced migration. In fact, between 1993-2010 49% did not. Therefore, I need to further subset the population of civil war-torn states to include only those with active refugee and/or IDP flows. This stage of selection may bias estimates. Here I employ the two-stage heckit correction method developed by Heckman (1979). The heckit method entails identifying a “selection equation” –a probit model that estimates the likelihood that a given civil war-torn state experiences forced migration. From the selection equation I obtain the inverse of the Mill’s ratio, which is the ratio of the probability density function to the cumulative distribution function of a distribution, and use it as a regressor in the outcome equation –a linear model estimated in ordinary least squares.

Given the nested nature of the panel data, I also employ robust standard errors clustered by country. Moreover, I include region fixed effects to control for unobserved region-to-region heterogeneity in estimating the first stage of the equation (selection into civil war-torn states). Finally, to control for temporal effects, I include a lag of the dependent variable in both stages of estimation. What I end up with is a two-stage heckit selection model with 361 civil war-torn country-years between 1993-2011. The following section formally introduces the model and estimation technique.

4.5.2 Estimation Technique

As mentioned in the previous section, my analysis involves a two-step estimation process. Given a population of war-torn states, I must first estimate the likelihood that such states produce refugee and/or IDP flows. In the subsequent step, among the remaining pool of states (civil war-torn states that produce forced migrant flows), I must identify factors that influence the share of forced migrants that are refugees. The first stage of analysis involves a basic probit regression written as follows,

$$P(D_{it} = 1 | Z_{it}) = \phi(Z_{it}\gamma) \quad (4.11)$$

where D_{it} indicates forced migration ($D_{it} = 1$ if country i experienced forced migration in time t and $D_{it} = 0$ otherwise), Z_{it} is a vector of explanatory variables, γ is a vector of unknown parameters, and ϕ is the cumulative distribution function of the standard normal distribution. In the second stage I include a transformation of the predicted individual probabilities (inverse Mill’s ratio) as an additional regressor in a model I estimate using OLS, which is notated as,

migrants does. Because selection at the first stage poses no threats of inference, I only present results of the standard two-stage heckit model. Please see Appendix C for the results of the multi-stage selection model.

$$f_{it}^* = X_{it}\beta + u_{it} \quad (4.12)$$

where f_{it}^* denotes the ratio of refugee flows to total migrant flows in country i in time t , which is not observed if the country does produce any forced migrants in a given year. Based on equations 4.11 and 4.12, the conditional expectation of the proportion of refugees to all migrant flows given the country experienced forced migration is written as follows,

$$E[f_{it}|X_{it}, D_{it} = 1] = X_{it}\beta + E[u_{it}|X_{it}, D_{it} = 1] \quad (4.13)$$

If we assume the error terms are jointly normal (i.e. multivariate normal distribution) then we obtain the following,

$$E[f_{it}|X_{it}, D_{it} = 1] = X_{it}\beta + \rho\sigma_u\lambda(Z_{it}\gamma) \quad (4.14)$$

where ρ is the correlation between unobserved determinants of propensity to produce forced migrants and unobserved determinants of the ratio of refugee flows to total forced migrant flows u , σ_u is the standard deviation of u and λ is the inverse Mill's ratio evaluated at $Z_{it}\gamma$. Hall (2002) argues that the standard two-step estimator results in inconsistent standard error estimates. This can be overcome using a variety of robust methods.

Alternatively, rather than estimating the equations using the standard two-step process, a maximum likelihood estimator (MLE) approach can be used.⁶ While the two-step method *controls* for the effect of variables in the selection equation on the outcome equation by including the inverse Mill's ratio, the MLE approach removes the effect of the variables in the selection equation from the outcome equation altogether.⁷ I present the findings of both in the results section.

A final note in regards to the estimation technique must be made before moving on to the data. For the heckman correction to successfully remove bias three key assumptions must be met. First, the standard estimation assumptions of both the outcome and selection equation equations must not be violated. Second, the selection equation must be specified well. Third, and perhaps most difficult to meet, at least one significant variable must affect the selection equation but have no independent significant effect on the outcome equation. In other words, one or more variables in the selection equation must act as an *instrument* that affects the

⁶This analysis was conducted in [R's sampleSelection](#) package, which uses the Newton-Raphson algorithm by default to maximize the log-likelihood function of the estimator. Alternative algorithms produced near identical results.

⁷See [Hall \(2002\)](#) for more.

probability of a country experiencing forced migration in a given year but not the composition of that migration (i.e. ratio of refugees to total forced migrants).

I posit that the effect of ethnic civil wars on the composition of forced migrants is conditional on the presence of one-sided violence. One-sided violence during nonethnic civil wars should either have no effect on the composition of forced migrants or should reduce the share of refugees relative to total forced migrants (i.e. increase the share of IDPs). Similarly, ethnic conflict, absent, one-sided violence, should have no effect on the composition of forced migrants. However, I also have argued that civilians respond to concerns over their personal safety by relocating to safer regions either in or out of their countries. If one-sided violence not only predicts forced migration but also influences the composition of forced migration then it cannot be used as an instrument. On the other hand, if there exists no independent marginal effect of one-sided violence on the composition of forced migrants then the instrument is valid.

4.5.3 Data

First it must be noted that the population of cases under examination only include country-years in which at least one civil war occurs as defined by [Bartusevičius \(2016\)](#) newly released dataset on Categoricaly Disaggregated Civil Wars (CDC).⁸ Given certain data availability issues discussed below, the population of cases is restricted to between 1993-2009, which results in 365 country-year observations. [Figure 4.1](#) is a choropleth map of the 45 states ravaged by civil war during the sample time period.

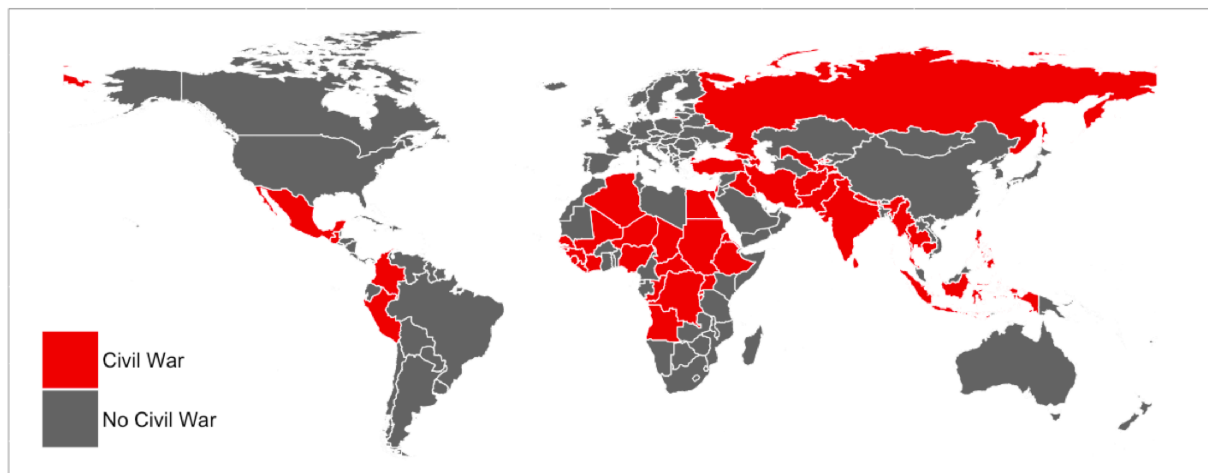


FIGURE 4.1: Conflict States 1993-2009

The dependent variable in both the selection and outcome equations is drawn from the [UNHCR](#)'s database of persons of concern.⁹ This database includes complete records of the countries

⁸<https://henrikasbartusevicius.com/cdc-1-0/>

⁹<http://www.unhcr.org/pages/4a013eb06.html>

of origin and asylum of refugees between 1951-2013 (incomplete data for 2014-2015). It also includes data on the number of internally displaced persons (and “persons in IDP-like situations”) between 1993-2013 (incomplete data for 2014-2015). Data for both refugees and IDPs are total stocks of migrants; however, the logic of population control is concerned with active migrant flows. To obtain flows, I take the first difference between the stock of migrants in time t and the stock of migrants in time $t - 1$ and then I truncate all negative values to zero.¹⁰

For the selection equation, where I estimate the probability of forced migration among civil war states, I simply add up the total number of refugee flows, IDP flows, and flows of persons in IDP-like situations and then convert this total into a binary measure; 0 for states with no active forms of forced migration and 1 for states with at least some form of forced migration (active refugee and/or IDP flows). For the outcome equation I am interested in the share of migrants that are fleeing refugees. To obtain this information I simply take the ratio of refugee flows to the total number of migrant flows. If Y_{it} is the dependent variable at time t in country i , then let R_{it} represent the number of refugee flows at time t in country i and I_{it} the number of IDP flows at time t in country i ,

$$Y_{it} = \frac{R_{it}}{R_{it} + I_{it}} \quad (4.15)$$

Prior to introducing the data of the regressors for each equation, a note must be made in regards to the operationalization of some of these covariates. As a check for robustness, whenever a data source is a count, such as the number of individuals displaced by natural disasters, I operationalize it in two ways, either as a natural log of the original count variable or a binary variable. To obtain binary scores, I simply choose zero as a threshold. For example, to operationalize the number of individuals displaced by natural disasters as a binary variable I code all country-years with at least one individual affected by natural disasters as a 1 and all remaining country-years (i.e. those with no individuals displaced) as a 0. For the logged counts, I just add one to the base and take the natural log. [Table 4.1](#) shows the raw values for each of these variables, [Table 4.2](#) shows their logged values, and [Table 4.3](#) shows their binary values.

The primary independent variable of interest (i.e. the instrument) in the selection equation is one-sided violence, which I draw from UCDP ([Sundberg 2009](#)). One-sided violence is defined as “the use of armed force by the government of a state or by a formally organized group against civilians, which results in at least 25 deaths” in a single event (see UCDP one-sided violence [codebook](#)).

¹⁰This may underestimate total migrant flows because some individuals may repatriate back to their homes, thereby reducing the total stock of forced migrants, which would reduce my measure of total migrant flows even if the total number of forced migrants fleeing conflict did not decrease. In effect, this measure conflates repatriation with decreases in refugee outflows. Because my population of cases is drawn from civil war-torn states I suspect the number of forced migrants returning home is likely very minuscule.

TABLE 4.1: Raw Values of Selection Equation Variables

Variable	N	Mean	St. Dev.	Min	Max
One-sided Violence	361	3,184.878	37,262.760	0	501,069
Forced Migrant Flows	365	70,796.720	248,107.600	0	2,107,111
Intercommunal Violence	361	55.413	199.246	0	2,127
Natural Disasters	361	373,037.400	733,338.500	32	4,695,110

TABLE 4.2: Binary Values of Selection Equation Variables

Variable	N	Mean	St. Dev.	Min	Max
One-sided Violence	365	0.660	0.474	0	1
Forced Migration	365	0.638	0.481	0	1
Intercommunal Violence	365	0.195	0.396	0	1
Natural Disasters	365	0.321	0.467	0	1

TABLE 4.3: Log Count Values of Selection Equation Variables

Variable	N	Mean	St. Dev.	Min	Max
log(One-sided Deaths)	361	3.525	2.832	0.000	13.125
log(Intercommunal Deaths)	361	0.974	2.042	0.000	7.663
log(Natural Disasters)	361	10.833	2.548	3.497	15.362

I operationalize this as the natural log of the number of deaths after having added 1 to the base and as a binary variable where 0 indicates less than 25 civilians deaths and 1 more than 25 (see [Table 4.2](#) and [4.3](#)). I also include a set of control variables. These include the natural log of number of deaths associated with intercommunal violence (and its binary operationalization, see [Table 4.2](#) and [4.3](#)), and the natural log of the number of individuals affected (killed and displaced) by natural disasters derived from The International Emergency Disasters Database [EMDAT](#) (and its binary operationalization, see [Table 4.2](#) and [4.3](#)).

I also include a set of variables that need no transformation. I control for involvement in international war by including a binary indicator of international conflict drawn from the Correlates of War [COW](#) dataset ([Sarkees and Wayman 2010](#)). I also suspect that forced migration is a feature more common to earlier stages of conflict than later ones as; thus, I control for conflict duration operationalized as the number of years since the civil war began. I also control for civil war intensity with an indicator coded as 1 if the civil war resulted in more than 1000 deaths in a given year and 0 if between 25-1000 casualties were recorded (obtained from the CDC dataset itself). Finally, I include a first order temporal lag of the DV (forced migration in time $t - 1$). [Table 4.4](#) shows the descriptive statics for each of these remaining variables.

TABLE 4.4: Untransformed Variables in Selection Equation

Variable	N	Mean	St. Dev.	Min	Max
Interstate War	365	0.055	0.228	0	1
Conflict Duration	365	5.638	4.640	0	17
Conflict Intensity	365	0.159	0.366	0	1

My primary independent variable of interest in this study, the *type* of civil war fought, appears in my outcome equation. Specifically, I am interested in whether countries experiencing ethnic civil wars produce different patterns of forced migration than countries experiencing nonethnic civil wars. I rely on [Bartusevičius \(2016\)](#) newly released CDC dataset. I opt for the CDC data over the Ethnic Armed Conflict (EAC) dataset and the Armed Conflict Data to Ethnic Power Relations (ACD2EPR) dataset because the EAC and ACD2EPR definitions of ethnic conflict are more restrictive than the theory tested in this study necessitates. EAC and ACD2EPR code conflicts as ethnic if two conditions are met; one, combatants “explicitly pursue ethno-nationalist aims, motivations and interests” and, two, the combatants “recruit fighters and forge alliances on the basis of ethnic affiliations” ([Cederman, Wimmer and Min 2010](#)).

The CDC, in contrasts, codes a conflict as ethnic if and only if its participating groups recruit members along ethnic lines. Because we cannot observe the aims and goals of ethnic groups (at best we can take their public announcements at face value) and because the logic of population control is concerned with observable markers of ethnicity and not with war aims, the CDC dataset is better suited to this study. If combatants recruit based on ethnic affiliation then it makes more sense that they will also target civilians based on ethnic affiliation. This is a simple binary indicator, 1 if a country-year experiences ethnic conflict and 0 if it experiences nonethnic conflict. However, some states experience both ethnic and nonethnic conflicts in the same year so to manage this overlap I code any country-year with at least one ethnic conflict as a 1 (i.e. as ethnic conflict).

An additional observable implication of the theory of the logic of population control is one-sided violence is an intervening variable between civil war type and composition of forced migration. Therefore, an interaction of ethnic conflict with one-sided violence should show an increase in the share of forced migrants that are refugees relative to the interplay between nonethnic conflict and one-sided violence. In other words, not only do I suspect that one-sided violence against civilians increases the likelihood of forced migration (stage 1, selection equation), I also suspect it increases the share of refugees relative to IDPs but *only* under conditions of ethnic conflict. I operationalize this conditional effect as an interaction term between one-sided violence and ethnic conflict in the outcome equation. Specifically, I interact the binary variable for ethnic

conflict with a binary variable indicating the presence of at least one event of civilian targeting.¹¹ All control variables are lagged one year.

In addition to the primary explanatory variables of conflict type and one-sided violence, I introduce a host of controls in the outcome equation as well, which can be classified into one of two groups – *neighborhood* and *domestic* factors. Neighborhood factors, on the one hand, refer to attributes of a country’s region that render seeking refugee across international borders more or less attractive. I expect civilians to be disinclined to seek refugee across international borders if they suspect the likelihood of being targeted is as high or higher in neighboring states as it is in his/her own country. Thus, I account for characteristics of neighboring states, which include controls for whether any neighbors of a country are experiencing civil wars (CDC dataset), one-sided violence (one-sided conflict UCDP dataset), and intercommunal conflict (non-state conflict UCDP dataset, see [Sundberg, Eck and Kreutz \(2012\)](#)). I also control for borders. I expect island countries to produce fewer refugees on average so I control for whether a country is an island or not. I also suspect fewer borders increases the burden neighbors face in accommodating fleeing refugees and, therefore, should decrease their willingness to accept large influxes of refugees. Thus, I control for the number of international borders a country shares with its neighbors.

Domestic factors, on the other hand, refer to characteristics of a country that make internal displacement more or less attractive for forced migrants. For example, I control for population density in case higher density countries make internal relocation more difficult and external migration more feasible. I also include a control for regime type using the Polity IV data ([Marshall and Gurr 2014](#)). All control variables are lagged one year. In addition to these controls, I also include temporal controls in both the selection and outcome equations. In both cases, I employ a first order temporal lag of the dependent variable to control for any autocorrelation in the errors. [Table 4.5](#) shows the descriptive statistics for each of these variables. The next section presents the findings.

TABLE 4.5: Variables for Outcome Equation

Variable	N	Mean	St. Dev.	Min	Max
Density	361	114.322	115.298	5.379	408.377
Island	365	0.085	0.279	0	1
Neighboring Civil War	365	0.597	0.491	0	1
Neighboring Intercommunal Violence	365	0.405	0.492	0	1
Neighboring One-sided Violence	365	0.699	0.459	0	1
Number of Borders	365	4.984	2.888	0	14
Refugee:Total	365	0.493	0.492	0.000	1.000
Ethnic Civil War	365	0.647	0.479	0	1

¹¹I also test this with a continuous operationalization of one-sided violence as discussed earlier.

4.6 Analysis

4.6.1 Organization of Results

I have estimated a number of Heckman style models using two different estimation techniques and two different operationalizations of the control variables; the results of only some of these models are presented in this section. One set of models is estimated using the traditional two-step method and the other set is estimated using MLE (Newton-Raphson algorithm). Furthermore, one set of models includes binary operationalizations of the control variables and the second uses the natural log of raw counts. What I end up with is 1) a model with dummy controls estimated using the two-step method, 2) model with dummy controls estimated using MLE, 3) a model with log counts estimated using the two-step method, and 4) another model with log counts estimated using MLE. Furthermore, each of these models contains two stages of estimation –a selection and outcome equation. I also estimate models with and without the interaction terms.

Including the two separate equations for each model, a total of sixteen equations are estimated. Instead of reporting the results for all sixteen models, I present the findings for the binary response variable models only because the results are robust across the two different operationalizations. The results for the model using log count explanatory variables can be found in Appendix C. The results are also largely robust to the two different estimation techniques; however, the findings of the two-step method and MLE did diverge in a couple of noticeable ways. Although the literature on selection models suggests MLE estimates are more consistent and robust (Hall 2002), I present results for both.

TABLE 4.6: Organization of Model Results

Model	Variable	Estimation	Effects	Section
Model 1	Binary	2step	Marginal	Results
Model 2	Binary	2step	Conditional	Results
Model 3	Binary	MLE	Marginal	Results
Model 4	Binary	MLE	Conditional	Results
Model 5	logCount	2step	Marginal	Appendix
Model 6	logCount	2step	Conditional	Appendix
Model 7	logCount	MLE	Marginal	Appendix
Model 8	logCount	MLE	Conditional	Appendix
Three-stage	Binary	2step	Marginal	Appendix

Table 4.6 lays out the components of each model and identifies whether its findings are located in the results section or the Appendix. The models shown in that table include the full two-stage model (i.e. not its selection and outcome components). Counting the outcome and selection

models for each model in results in 19 separate equations (the three-stage model has three equations). Next I turn my attention to the results of the eight equations of interests –selection and outcome stage models with and without interaction terms estimated using MLE and 2step estimation and a set a binary control variables (the first four models in bold found in [Table 4.6](#)).

4.6.2 Results

I begin by examining the results of the marginal effects selection model (i.e. without interaction effects) in the first stage of estimation. In this stage of estimation, I have regressed the probability of forced migration on the presence of international conflict, intercommunal conflict, one-sided violence, natural disasters, and civil war intensity. Model 1_{MLE} and Model 2_{2step} of [Table 4.7](#) shows the selection results of the ML and 2step estimation respectively for the model without interaction effects in its outcome stage. The positive sign and statistical significance of the coefficients of one-sided violence suggest that civilians respond to the risk of persecution by fleeing, either as refugees or IDPs, which confirms H₁. Episodes of forced migration in the previous year also increase the chances of forced migration, suggesting that 1) similar factors likely persist over the course of a conflict that push civilians to flee and 2) civilians learn from the past to inform their decisions of the future (if they see others persecuted and fleeing in time $t - 1$ they will be more like to flee themselves in time t). Forced migration is no more or less likely to occur earlier in conflicts, while occurrence of natural disasters, intercommunal violence, international conflict, and war intensity also all fail to reach statistical significance. It appears as though the number one factor that influences forced migration is the deliberate and organized targeting of civilians.

Now I turn to the outcome equation. ρ and σ in [Model 1](#) are both significant, suggesting that selection bias is present absent a well-specified selection equation in the first stage. In other words, Model 1's outcome equation confirms that using a two-stage Heckman like selection model is the appropriate approach –a finding that is also confirmed by Model 2's outcome equation, whose inverse Mill's ratio is also significant.

A number of factors affect the composition of forced migrants. For example, island countries produce significantly more IDPs as a share of total forced migrants than countries with territorial borders. More importantly, the nature of conflict divisions –whether a civil war is fought along ethnic lines or not –also affects the ratio of refugees to IDPs. The positive sign and statistical significance of the coefficients of the ethnic civil war variable suggests that states embroiled in ethnic civil war produce more refugees relative to total forced migrants than states embroiled in nonethnic civil war, which confirms H₃. This is true of the models estimated using both techniques (MLE and 2step). H₂ posits that although one-sided violence increases the chances

TABLE 4.7: Marginal Effects Models

	Model 1 _{MLE}		Model 2 _{2step}	
	Selection	Outcome	Selection	Outcome
Civil War Intensity	−0.10 (0.19)		−0.10 (0.20)	
Forced Migration _{t-1}	0.55** (0.15)		0.54** (0.15)	
Interstate War	0.24 (0.29)		0.23 (0.29)	
Conflict Duration	−0.02 (0.02)		−0.02 (0.02)	
Intercommunal Violence	0.03 (0.17)		0.05 (0.18)	
Natural Disasters	−0.19 (0.15)		−0.09 (0.15)	
Civilian Targeting	0.68** (0.15)	−0.12 (0.07)	0.69** (0.15)	−0.12 (0.07)
Ethnic War		0.18** (0.07)		0.15** (0.07)
Population Density		0.00 (0.00)		0.00 (0.00)
Island		−0.39** (0.13)		−0.37 (0.14)
Civil War _{Neighborhood}		0.04 (0.07)		0.04 (0.08)
Intcom.Viol _{Neighborhood}		−0.05 (0.06)		−0.05 (0.07)
Civil.Targ _{Neighborhood}		−0.05 (0.08)		−0.04 (0.08)
Borders		−0.02 (0.01)		−0.02 (0.01)
Ref:Forced Migrants _{t-1}		0.35** (0.08)		0.35** (0.08)
(Intercept)	−0.29 (0.16)	0.67** (0.12)	−0.22 (0.16)	0.02 (0.35)
σ		0.39**		
ρ		0.47**		
invMillsRatio				0.88**
Adj. R ²		0.17		0.14
Num. obs.	227	365	227	365

Note:

*p<0.05; **p<0.01

of forced migration (H_1), it has no independent marginal effect on the composition of forced migrants (i.e. the share that is composed of either IDPs or refugees). Both models fail to uncover any statistically significant relationship between one-sided violence the dependent variable of the outcome equation, which confirms H_2 . This is an extremely important finding because had such an effect existed, one-sided violence could not be used as an instrument in the selection equation, while the estimates of the outcome equation would remain biased (i.e. the selection bias would persist).

Surprisingly, there is no evidence to suggest any of the remaining neighborhood effects influence the share of forced migrants that are refugees. The number of international borders, the occurrence of civil war, intercommunal violence, or one-sided violence in neighboring states has no statistically significant effect on the composition of forced migrants. Although surprising, this may be accounted for by the fact civilians likely have incomplete information in regards to the likelihood of persecution abroad. Their impression of where the safest place to flee is likely formed by information they have about conditions at home and not those abroad. Therefore, we would civilians react to *domestic* factors rather than *neighborhood* ones when information is scarce -a typical feature of warfare. Regime type and population density also fail to reach significance.

I now turn to the results of the conditional effects model. Because the interaction effects are not introduced until the second stage, the results of the conditional effects selection equation are nearly identical to the results of the marginal effects selection equation. Here, I am only interested in the *conditional* effects of ethnic conflict and one-sided violence on the share of forced migrants that are refugees. To do this, I can examine the interaction between civil war type and one-sided violence. Table 4.8 shows the results of the MLE and 2step models, both of which identify a statistically significant relationship between the share of forced migrants that are refugees and the interaction between civil war type and one-sided violence, which confirm H_4 and H_5 .

Moreover, the marginal effect of ethnic civil war when one-sided violence does *not* occur is insignificant, which also confirms H_6 . The one difference that arises between the two estimation techniques, is that the marginal effect of one-sided violence on forced migrant composition is insignificant for nonethnic civil wars when estimated using the 2step method but significant and negative when estimated using the MLE approach. The MLE results suggest one-sided violence during nonethnic civil war increases the share of IDPs rather than the share of refugees. Although not shown in the coefficient tables, the marginal effect of one-sided violence on the composition of forced migrants during ethnic civil wars is negative and statistically significant. These findings suggest that civilians fleeing ethnic civil wars have more incentives to seek refugees across borders than in “safer” regions of their home countries if the violence they are fleeing is

TABLE 4.8: Conditional Effects Models

	Model 3 _{MLE}		Model 4 _{2step}	
	Selection	Outcome	Selection	Outcome
Civil War Intensity	−0.10 (0.19)		−0.10 (0.20)	
Forced Migration _{t-1}	0.53** (0.15)		0.41** (0.15)	
Interstate War	0.24 (0.29)		0.15 (0.30)	
Conflict Duration	−0.02 (0.15)		−0.02 (0.02)	
Intercommunal Violence	0.03 (0.18)		0.05 (0.18)	
Natural Disasters	−0.20 (0.15)		−0.09 (0.15)	
Civilian Targeting	0.68** (0.15)	−0.27** (0.09)	0.69** (0.15)	0.04 (0.18)
Civilian Target X Ethnic War		0.26** (0.12)		0.25** (0.08)
Ethnic War		0.18** (0.07)		−0.03 (0.05)
Population Density		0.00 (0.00)		0.00 (0.00)
Island		−0.41** (0.13)		−0.39 (0.14)
Civil War _{Neighborhood}		0.05 (0.07)		0.04 (0.14)
Intcom.Viol _{Neighborhood}		−0.05 (0.06)		−0.04 (0.06)
Civil.Targ _{Neighborhood}		−0.07 (0.08)		−0.06 (0.08)
Borders		−0.02 (0.01)		−0.02 (0.02)
Ref:Forced Migrants		0.25** (0.05)		0.34** (0.08)
(Intercept)	−0.29 (0.16)	0.79** (0.13)	−0.22 (0.16)	0.15 (0.35)
σ		0.39**		
ρ		0.50**		
invMillsRatio				0.89**
Adj. R ²		0.17		0.14
Num. obs.	227	365	227	365

Note:

*p<0.05; **p<0.01

specifically directed towards them. This has important implications for the spread of conflict, which I turn to in my concluding section.

4.7 Conclusion

In this chapter I have introduced a theory, which I label the *logic of population control*, that accounts for some of the variation in patterns of forced migration resulting from on going civil wars. The logic of population control draws on the findings of the previous chapters as well as the extant literature on ethnic conflict and civilian targeting.

In the previous chapters I argued that coethnic refugees contribute to increased levels of violence because they can alter the delicate *ethnic balance of power* in host states with volatile ethnic relations. Because refugees are a potential source of recruitment for coethnic groups in the host state, rival groups, rightly or wrongly, view them as a threat. It is for this very same reason that combatants also view local civilians perceived as loyal or sympathetic to their rivals as a threat. However, whether combatants target civilians loyal to their rivals largely depends on if they can determine whether the civilian is a supporter or not. The literature on civilian targeting suggests ethnic markers aid in this process of identification.

If ethnic markers increase violence against civilians, then civilians fleeing ethnic conflict are more constrained in their choice of destination. Simply relocating to regions away from the crossfire may not be enough; combatants may target the regions they flee to as well. Therefore, I have hypothesized that states suffering ethnic civil wars on average should produce more refugees relative to IDPs than states afflicted with nonethnic civil war.

To test this proposition I applied a two-step Heckit style selection model to a population of all country-years experiencing civil war between 1993-2010. The results of my analysis confirm my expectations that a statistically significant difference exists between patterns of forced migration resulting from ethnic and nonethnic civil wars.

The finding that ethnic conflicts contribute to refugee flight across borders has serious implications for the study of conflict contagion more generally. Previous literature has examined the conditions under which states produce more or less refugees, but these studies, with one or two notable exceptions, have failed to appreciate the counterfactual to refugee flight, which is internal displacement.

Civilians face two primary choices during conflict that can impact their very survival; 1) to stay put (either as a civilian or to take up arms) or to flee, and 2) whether to flee to other regions of their own countries or to seek asylum in foreign countries. The current literature on

forced migration either aggregates refugees and IDPs into a single category or focuses exclusively on refugee flows. Aggregating all forms of forced migration into a single category allows us to identify the factors that contribute to civilian flight, the first choice civilians in conflict make, but it cannot speak to the second choice they face -their destination of refugee. Moreover, examining only refugee flows leads researchers to under appreciate how IDP flows can also contribute to conflict contagion. For example, opposition groups in neighboring states may find safe zones for IDPs to be useful as safe havens for their own activities.¹² Thus, examining refugees and IDPs together but as analytically distinct elements can help match different mechanisms of conflict contagion with different patterns of forced migration.

¹²The Kurdish PKK in Turkey has found support in the safe zones along Syria's Turkish borders where many Kurdish IDPs have amassed. For more information on how the PKK has used these regions in Syria see <http://www.al-monitor.com/pulse/security/2016/04/turkey-pkk-clashes-last-stronghold.html>. For more on IDPs and the spread of conflict see [Bohnet \(2012\)](#).

Chapter 5

Conclusion

Key Contributions

Forced migration is both a consequence and a cause of the spread of conflict across borders. States have a vested interest in controlling the populations they rule and rebels have incentives to undermine this control. Under conditions of ethnic conflict, population control is facilitated by ethnic markers that aid combatants in identifying loyal populations and those sympathetic to rivals. The targeting of noncombatants, civilians and refugees alike, is a natural consequence of this friction. The aim of this study has been to examine this friction under a number of different conditions.

This study makes five main contributions to the literature on forced migration and conflict. First, it has identified a refugee-conflict nexus at the substate region in Lebanon. There exist no substate analyses of refugees and conflict outside of sub-Saharan Africa. Thus, this represents the first test of that association in a previously unexplored region of the world. Second, it identifies ethnic tensions as a conditioning factor between violence and the presence of refugees -regions with pre-existing ethnic tensions that host refugees are more likely to see violence than ethnically stable regions that host similar refugees. I argue this is the result of changes in the ethnic balance of power between competing groups. Third, this study has shown that ethnically mobilized civil wars are more likely to produce refugees relative to IDPs than nonethnic civil wars. Finally, the empirical analysis has shown that refugees do not actively contribute to conflict onset but, instead, shape the way conflict is waged. Existing studies on conflict and refugees typically highlight the role refugees play as active combatants -in the form of recruits or public advocates. However, this research has shown that the association between refugees and violence is largely confined to a particular form of violence, one-sided violence, which suggests refugees are victims of the spread of violence as much if not more than they are active

contributors to it, which has a number of significant policy implications that are discussed in the following section.

Policy Implications

The conventional wisdom in the literature on forced migration and conflict suggests refugees pose a danger to the host state, typically in the form of recruits and added pressure on the local population. This is most succinctly captured in the notion of the “refugee warrior“. Yet analyses of disaggregated features of conflict suggest refugees, far from being active participants in conflict, are typically its victims -even in the regions they have sought refuge. Moreover, this association is mediated by ethnic ties between forced migrants and their hosts, which has important policy implications.

In terms of policy, these findings may point to alternative means of organizing refugees in the host population. Currently, the UNHCR advocates for the establishment of camps to house refugees. However, if refugees are susceptible to targeting by rebel groups, these camps render refugees as vulnerable as sitting ducks. Perhaps a more effective way to organize the distribution of refugees in a host state is to diffuse their presence across different regions and to limit the number of refugees to regions with volatile ethnic tensions. This not only minimizes the added pressures local populations face, it also limits the ability of rival groups to target specific camps as a means of undermining the support of their rivals or bolstering their own recruitment efforts. There is evidence to suggest that at least some decision-makers have come to this same conclusion. For example, leaders in Lebanon, a state who has hosted numerous Palestinian refugee camps scattered across its landscape for decades, has refrained from establishing centralized camps for Syrian refugees and, instead, has opted for a policy of general dispersion. Despite the low-intensity violence that is currently rocking Lebanon (which has been well documented in this study), the country has managed to avoid the type of high-intensity violence that has become commonplace in Iraq and Turkey in recent years.

The findings of this analysis also suggest that growing concerns over the influx of refugees to Western states may be overblown. [Figure 5.1](#) shows Google searches for the terms “Refugees”, “Syrian Civil War”, and “Terrorism” since 2014 along with the actual incidents of conflict. The y-axis in the graph represents Google’s normalized search metric for each search term ¹. The x-axis represents the time range from 2014-2016, along with the locations of major terrorist incidents in the Western world during that same time period.

¹According to Google Trend’s each data point is divided by the total searches of the geography and time range it represents. The resulting numbers are then scaled to a range of 0 to 100. See the following link for more information <https://support.google.com/trends/answer/4365533?hl=en>

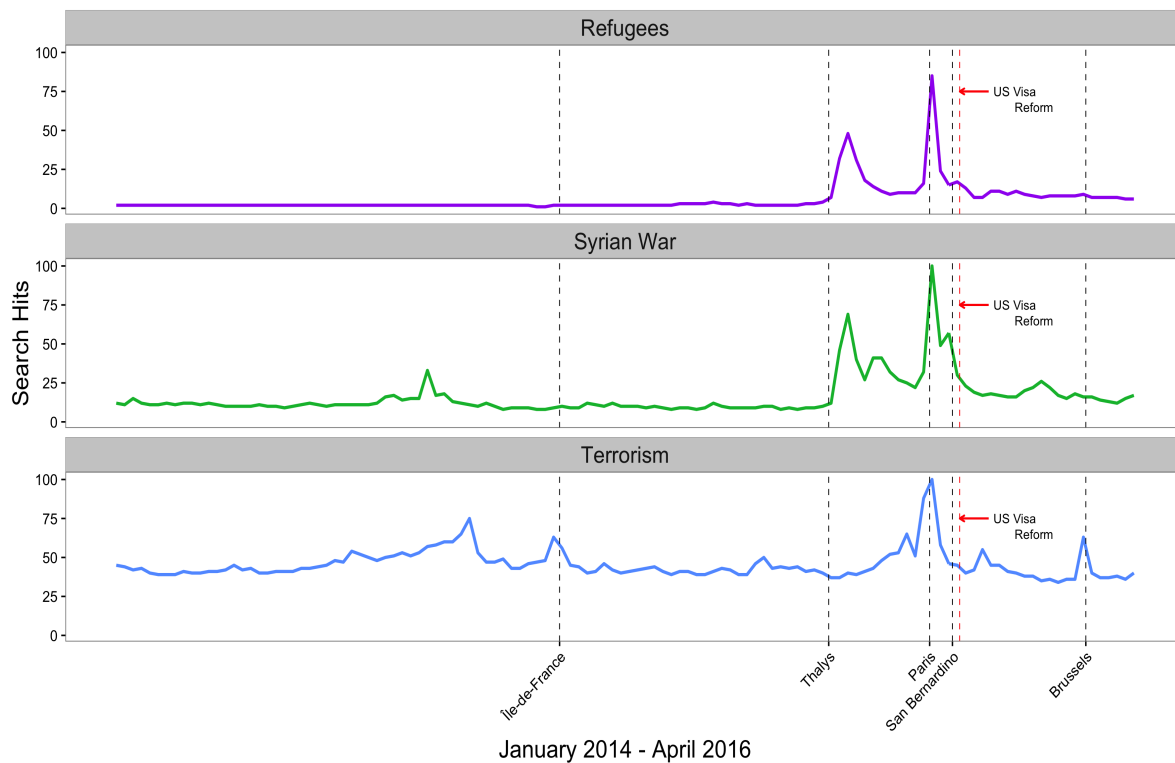


FIGURE 5.1: Google Keyword Search Hits

Figure 5.1 shows that searches for refugees, the Syrian civil war, and terrorism co-vary to a considerable degree, suggesting popular perceptions of refugees, war, and terrorism are tied. While it is no surprise that people associate refugees with war, the fact that terrorist incidents are closely followed by commensurate changes in popular search habits, represents a more general fear of refugees held by the public. The red vertical line identifies the US Congress's vote on bill H.R. 158 that successfully revoked visa waiver stipulations for European dual citizens of Sudan, Iran, Iraq, and Syria, in the wake of the San Bernardino attacks. Whether politicians who passed H.R. 158 were genuine in their concern for the link between refugees and terrorism or whether they were capitalizing on fears of the public, the popular perception of refugees as vectors of conflict transmission has some clear effects on public policy. Based on the findings of this study, I argue that we can do more to limit violence associated with refugees by doing more to guarantee their security. For as long societies wage war, people will be fleeing it.

Appendix A

Appendix: Codebook

Event data on substate conflict in Lebanon was manually coded from two local news wire services -Naharanet and National News Agency. Naharnet is Lebanon's first independent Digital Media providing real-time news in English.¹ National News Agency is Lebanon's official state news and information service run by the Ministry of Information.² The data collection efforts cover the period from March 2013 to April 2015. Each observation is coded for seventeen variables. Only events that are related to political conflict or sectarian conflict are coded for the dataset.

It must be noted, for the purposes of this study only certain categories of conflict are included in the sample. For example, events related to illicit behavior (e.g. drug trafficking and human smuggling undertaken by militants) that do not involve violence are dropped from the sample. Also dropped are all events related to cross-border conflict with other countries, such as Syrian and Israeli military raids into Lebanese territory as well as cross-border shelling of Lebanese towns and villages from the Syrian side of the border.

Actors, Definitions, and Variables

The variable name used in the dataset is found in brackets.

1. The **date** the event occurred [date]

- *Description:* Day, Month, and Year in which the event took place.
- *Notes:* If the event took place over more than one day. Each day was coded as a separate event.

¹<http://www.naharnet.com>

²<http://nna-leb.gov.lb/en>

2. The **actor's identity** [actor]

- *Notes:* The difference between actor and target is that actor refers to the participant who instigates the conflict event. Cases in which the identity of the initial instigator is unknown then the first participant named in the primary sourced is coded as the actor.
- *Description:* Actor identity refers to the specific identity of an actor involved. For example, if a Lebanese civilian attacks a Syrian refugee then the actor identity is coded as a Lebanese civilian. If Lebanese police arrest a militant then the actor's identity is Lebanese Internal Security Force.
- *Values:* The following are a list of all coded actors:
 - **lebauth:** Unspecified Lebanese authority
 - **lebmilint:** Lebanese Military Intelligence
 - **lebarmy:** Lebanese Armed Forces
 - **lebisf:** Lebanese Internal Security Forces
 - **lebmun:** Lebanese Municipal Authorities (includes local police, mayor's office, etc.)
 - **palisf:** Palestinian Internal Security Forces
 - **isrlmil:** Israeli Military
 - **syrmil:** Syrian Military
 - **syrctg:** Syrian Coast Guard
 - **lebcivil:** Lebanese civilian
 - **isrlcivil:** Israeli civilian
 - **syrref:** Syrian refugee or Syrian national not identified specifically as a rebel or with a rebel group
 - **palref:** Palestinian refugee or Palestinian national not identified specifically as a rebel or with a rebel group
 - **hezb:** Hezbollah militant group (shia)
 - **amal:** Amal militant group (shia)
 - **rebpal:** Palestinian rebel groups such as Popular Front General Command, Fatah, Hamas, etc.
 - **rebfsa:** Rebel group Free Syrian Army (sunni)
 - **rebisf:** Rebel group Islamic State , includes offshoots such as Abdullah Azzam Brigades (sunni)
 - **rebanf:** Rebel group Al Nusra Front (sunni)
 - **rebskh:** Sheikh Khaled Hablas rebel group (sunni)

- **rebbsaa**: Sheikh Ahmad Asir rebel group (sunni)
- **rebunid**: Unidentified rebel group
- **unifil**: United Nations Interim Force in Lebanon
- **unrwa**: United Nations Relief and Works Agency
- **other**: None of the above (i.e. national not from Lebanon or Syria)

3. The **actor's group** [actor_group]

- *Description*: Actor Group refers to the religious or institutional background of an actor. For example, a Sunni Lebanese civilian would be coded as a Sunni as would a Sunni militant. The institution of an actor is coded only if the actor is acting on behalf of a government. For example, a Lebanese Army force and a Lebanese Police would both be coded as Lebanese government.
- *Values*: The following are a list of all coded groups:
 - **sunni**: Sunni Muslim
 - **shia**: Shia Muslim
 - **druze**: Druze
 - **alawi**: Alawi
 - **jewish**: Jewish
 - **christian**: Maronite, Catholic, Protestant, Armenian Catholic, Greek Orthodox
 - **palgov**: Palestenian Authority
 - **syrgov**: Syrian Government
 - **lebgov**: Lebanese Government
 - **isrlgov**: Israeli Government
 - **intlorg**: International Organization (such as UNRWA or UNHCR)
 - **other**: None of the above

4. The **target's identity** [target]

- *Description*: The difference between actor and target is that actor refers to the participant who instigates the conflict event. Cases in which the identity of the initial instigator is unknown then the first participant named in the primary sourced is coded as the actor.
- *Values*: See actor identity

5. The **target's group** [target_group]

- *Description*: The difference between actor and target is that actor refers to the participant who instigates the conflict event. Cases in which the identity of the

initial instigator is unknown then the first participant named in the primary sourced is coded as the actor.

- *Values:* See actor group

6. The number **killed** [dead]

- *Description:* The number of individuals, civilians and non-participants included, who were killed during the course of the event or as a result of it)
- The coding for number killed is not always perfectly reflected in the source URL for two reasons 1) the numbers are updated as new news articles are reported on an event (even though those other URLs are not coded in the source column) and 2) because some events were coded using additional URLs that are not reported,

7. The number **injured** [injury]

- *Description:* The number of individuals, civilians and non-participants included, who were injured during the course of the event.

8. The number **arrested** [arrest]

- *Description:* The number of individuals who were arrested during the course of the event. Arrests only refer to arrests made by Lebanese government.

9. The **category** of the event [category]

- *Description:* This variable categorizes each event according to the type of conflict event, ranging from clashes involving numerous people on both sides to assassinations, bombings, and even hostage taking.
- *Values:* Many different categorizations
 - **execut:** Event involves execution outside of battles unrelated to basic shooting (such as execution of hostages)
 - **arrestraid:** Event involves raid or arrest by military or internal security forces
 - **bombard:** Event involves bombardment with weapons such as surface-to-air grade missiles outside the context of a clash (defined below)
 - **assault:** Event involves use of violence b/w 1-3 individuals (not including ‘shooting’), involves weapons other than projectiles and guns.
 - **shooting:** Event involves shooting arising from dispute b/w 1- 3 individuals (such as individual murder)
 - **protest:** Event involves protest or other collective actions such as solidarity marches and strikes but also includes acts such as rioting and forced closures of streets and shops

- **abduction**: Event involves use of kidnapping, abduction, ransom, or hostage-taking
- **clashes**: Event involves clashes bw armed groups (4+ individuals involved in total) with guns, artillery, grenades, rockets, missiles, and/or mortars that wholly takes place within Lebanon (includes siege)
- **suicbatt**: Event involves suicide bombing
- **bombatt**: Event involves explosive device besides suicide attack within Lebanon (IED, bomb, intentional mine, not hand grenades unless placed as a detonated bomb -so not used with ‘hands’ as a projectile weapon such as a missile launcher, etc.). ”Non-events” are not coded. For example, if a bomb is found that has not gone off, that is not coded. is NOT coded. Only events that transpire are coded
- **curfew**: Event involves imposition of curfew on refugees by authorities at any level of government
- **crossclash**: Event involves land fire (gunfire, mortars, missiles but not bombs from planes) from across Syrian or Israeli border, or it involves territorial violation by Israeli or Syrian governments
- **aerbom**: Event involves aerial bombing (cross-border included)
- **airspace**: Event involves nonviolent illegal entry of Lebanese airspace (Syria and Israel are the usual actors)
- **maritime**: Event involves nonviolent illegal entry of Lebanese territorial waters
- **territ**: Event involves nonviolent illegal entry of Lebanese land territory

10. The **classification** of the event [classification]

- *Description*: This variable classifies events according to the nature of the underlying tension between each actor-target.
- *Values*: Seven classifications
 - **perscon**: Event associated with conflict over individual or peronal act of violence/illicit activity (i.e. non political/economic/social such as family dispute)
 - **relcon**: Event associated with conflict bw religious groups and/or social groups (includes arrest of Syrian refugees for lack of refugee papers)
 - **polcon**: Event associated with conflict over policy between competing political factions or citizens and government (all incidents of terrorism or related to terrorist groups coded as polcon including arrests for possession of military style weapons)
 - **illcon**: Event associated with conflict over illicit group activities associated with militants (such trafficking, drugs, etc.)

- **syrcrosscon**: Event associated with cross border conflict from Syria (includes illegal entry into the country by refugees, fighters, etc.). If a single cross-border assault (such as aerial bombing or mortar/gunfire) hits multiple villages within Lebanon, each location that is struck is coded as a separate event
 - **isrcrosscon**: Event associated with cross border conflict from Israel
 - **econcon**: Event associated with conflict over economic issues (for example protest over food prices, these are very rare in the data simply because news reports rarely report any information that a coder can infer to suggest the event involves economic related issues/ demands/tensions).
11. The **town** the event took place [town]
 - *Notes*: Closest town to the event.
 12. The **region** the event took place [region]
 - *Description*: This refers to informal regional names.
 13. The **district** the event took place [district]
 14. The **governate** the event took place [governate]
 15. The **latitude** of the event [latitude]
 - *Notes*: Closest geo-coded point of reference to the event
 16. The **longitude** of the event [longitude]
 - *Notes*: Closest geo-coded point of reference to the event
 17. The **URL** of the primary sourced used to code the event [url]
 - *Notes*: When multiple sources identify the same event only one source is listed.

General Notes

1. If a single cross-border assault (such as aerial bombing or mortar/gunfire) hits multiple villages within Lebanon, each location that is struck is coded as a separate event
2. Arrests only refer to arrests made by the Lebanese government (and not for example by Syrian military forces who enter into Lebanon to detain militants. That would be coded as a '**syrcrosscon**' (illegal Syrian entry of Lebanese territory).

3. Only the primary two agents are coded. For example if a clash breaks out between Sunni and Shia groups and as a direct result of those clashes the army arrests one or more of the actors/targets involved then only the two actors/targets that were initially involved in event are coded even though the event includes an arrest by an army officer. This is only the case if the arrest happens immediately within the time frame of the event (same day for example, anything past a day gets coded as a separate event)
4. Events cannot last more than one day -for events related to one another in a larger battle, each day is coded as an independent event.
5. If an event involves action by more than one participant (i.e. a reaction) the event is coded according to the attributes of the 1st action. For example, if an event involves an abduction or hostage taking by actor A and the victim is target A but the news wire also states that the hostage was freed on the same day after the army arrested actor A then only the abduction as coded for the variable '**category**' (and not the subsequent arrest). Similarly, the actor is coded as the first actor and the target is coded as that first actor's target (even though eventually our actor A in the example above moves from being an actor to a target when he is arrested by the army). Keep in mind, however, because the event ultimately resulted in the arrest of actor A the variable '**arrest**' is still coded as 1 (or if he was killed in the raid then 1 for dead).
6. 'Non-incidents' are not included such as 'Lebanese Amry Intelligence discovered an detonated bomb in a parked car'
7. Explosions caused by land mines from previous conflict are not included (unfortunately more common than one would think)
8. If two different event classifications occur at once (for example clashes occur between two families and security forces get involved and arrest those involved then the very first action to perpetuate the conflict is coded. In this case if arrests were made immediately at the time of the clashes the event would be coded as '**clashes**' and not '**arrestraid**' despite the fact that an arrest was made *and* the variable for arrest would still be coded as 1. Conversely if an arrest was being made and clashes erupted or shooting erupted immediately (within the frame of action such as during the time the arrest was being made or immediately following it within a couple hour block) as a result of the attempted or successful arrest then the event would be coded as an '**arrestraid**' and not as '**clashes**' or '**shooting**'.
9. It is not always clear who the actor and who the target are. Usually the actor is the agent who initiates the confrontational event but if the news report does not include that info or if no clear initiator even existed then the default is to always use civilians as the target.

In cases where civilians are not one of the agents then it reverts to the government. For example, if an unidentified rebel group (**rebunid**) is in a clash with Lebanese army (**lebarmy**) and it's unclear who the target and who the actor is then the actor by default gets set as the government. In cases in which neither civilians nor Lebanese government actors are involved then the group that sustains most injury or death gets coded as the actor and if it's still unclear then the first group in the news wire to be mentioned is coded as the actor.

10. Incidental victims do not get coded as targets though if they are injured or killed a death and/or injury will be coded. For example, if a clash between the army and some unidentified rebel group takes place and a Palestinian refugee is killed in the cross fire then the target and actors would be either Lebanese army or the rebel group (depending on who initiated) but it would not be the Palestinian refugee.
11. Only drug trafficking arrests are reported, so no events involving drug possession or drug transport are coded unless violence is involved in which case it is reported even if it is just possession or transport
12. Only violent events are (with the exception of nonviolent arrests of militants and nonviolent illegal entry of militants into the country) that are related to political conflict are included

Appendix B

Appendix: Data Visualization

FIGURE B.1: Refugee Flows and Violence 2013-2015

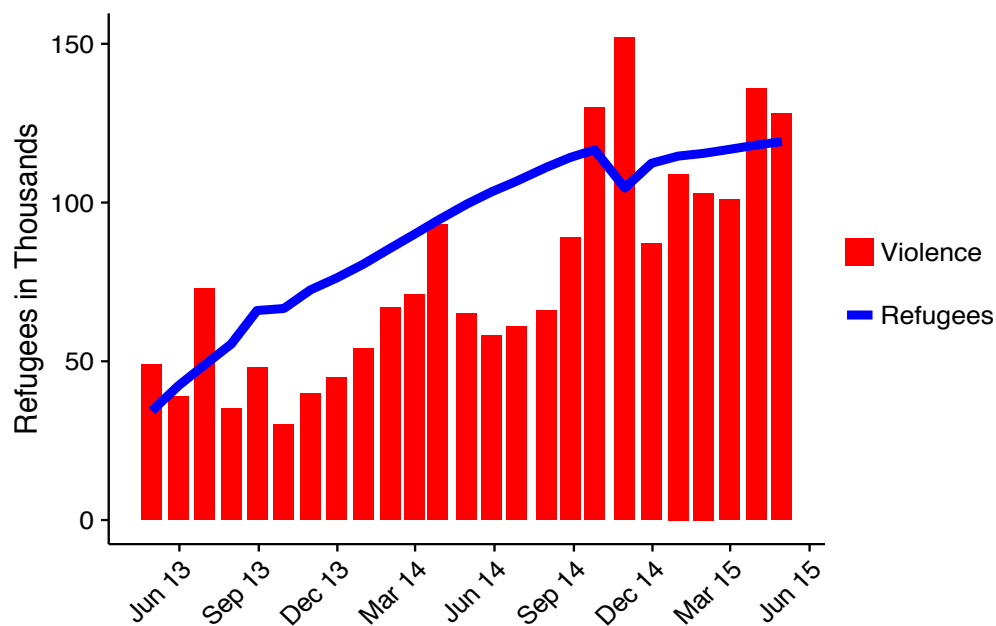


FIGURE B.2: Violence Severity 2013-2015

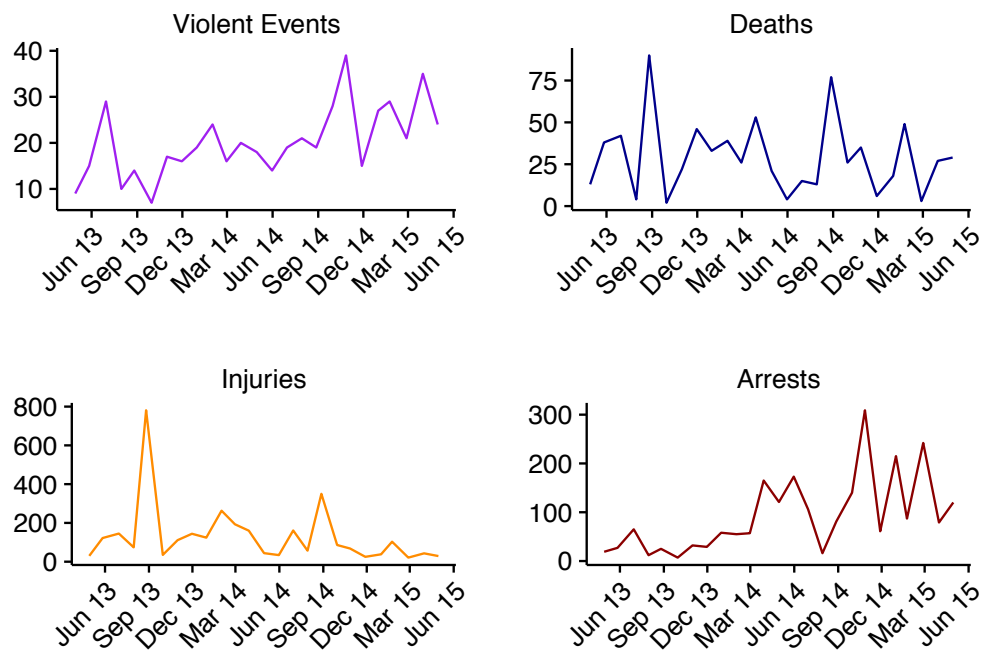


FIGURE B.3: Violence by District 2013-2015

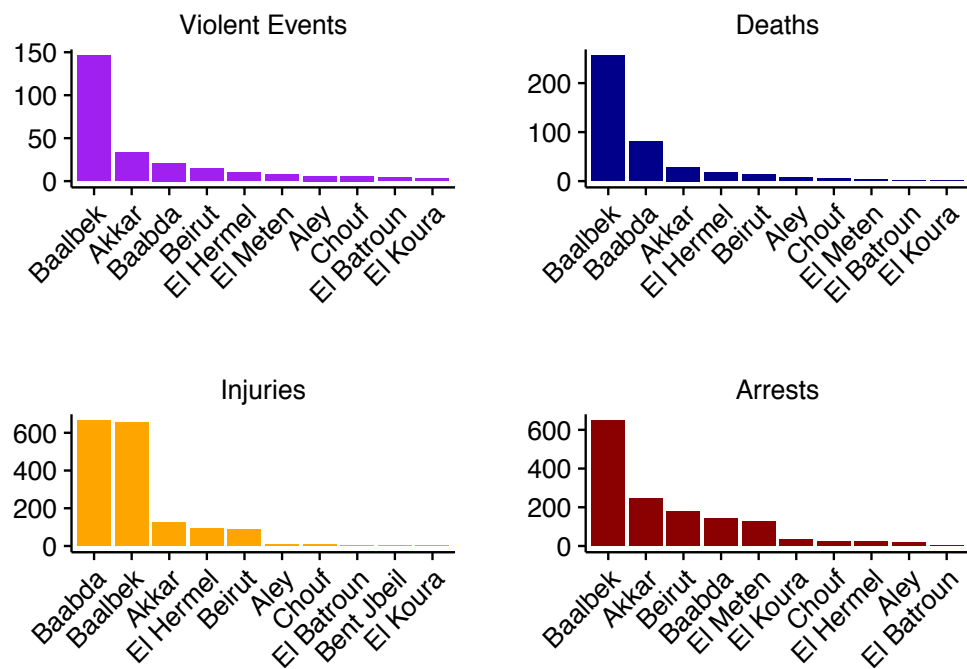


FIGURE B.4: Syrian Refugees as a Share of Population 2013-2015

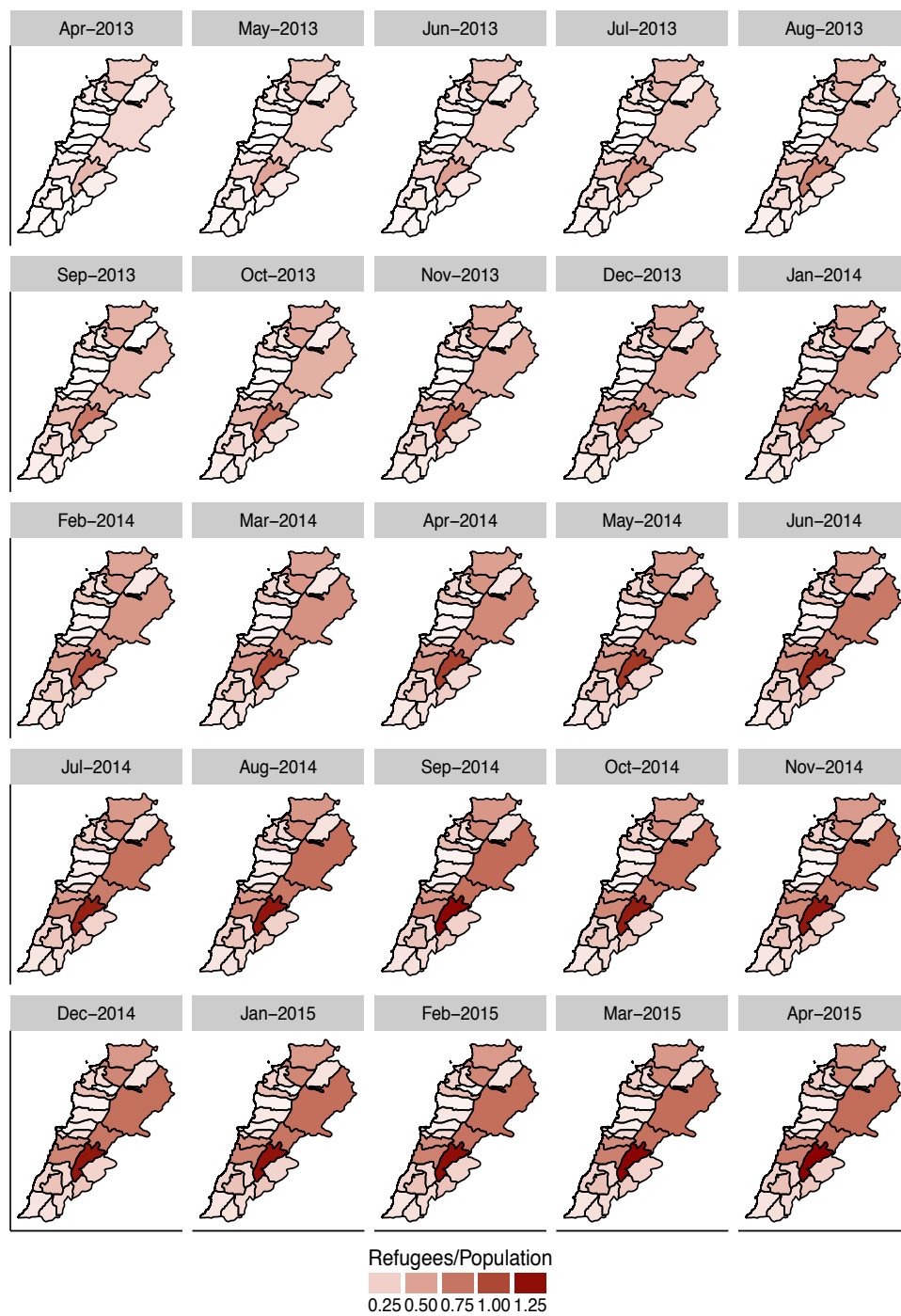


FIGURE B.5: Ethnoreligious Tapestry of Lebanon

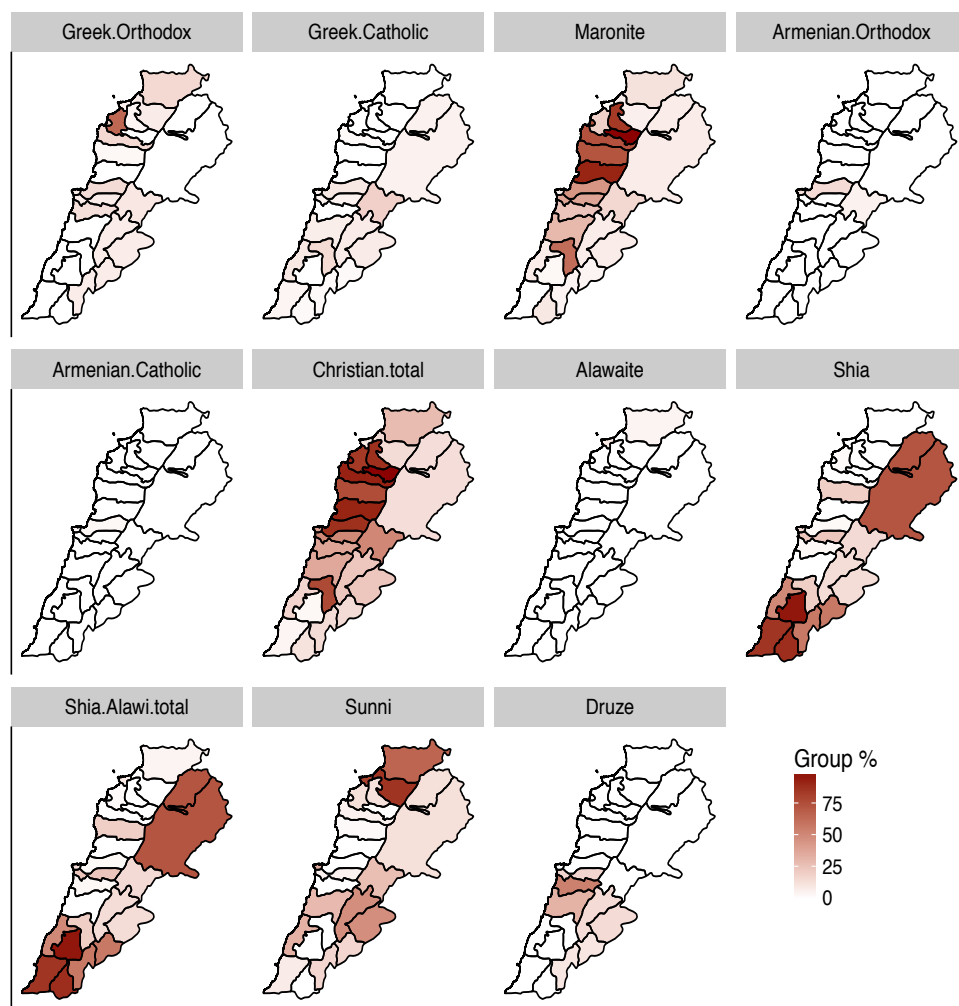
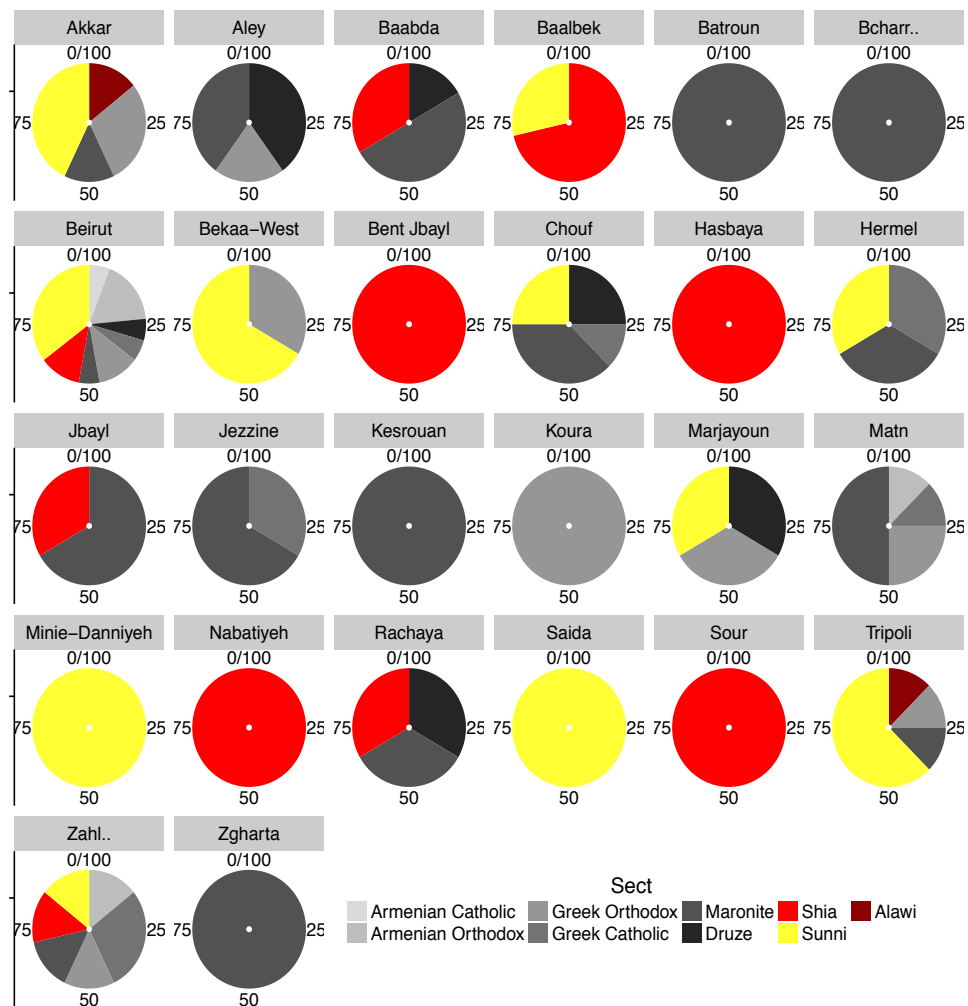


FIGURE B.6: Ethnoreligious Groups by District



Appendix C

Appendix: Log Count Results

hello

TABLE C.1: Organization of Model Results

Model	Variable	Estimation	Effects	Section
Model 1	Binary	2step	Marginal	Results
Model 2	Binary	2step	Conditional	Results
Model 3	Binary	MLE	Marginal	Results
Model 4	Binary	MLE	Conditional	Results
Model 5	logCount	2step	Marginal	Appendix
Model 6	logCount	2step	Conditional	Appendix
Model 7	logCount	MLE	Marginal	Appendix
Model 8	logCount	MLE	Conditional	Appendix
Three-stage	Binary	2step	Marginal	Appendix

TABLE C.2: Marginal Effects Models with Log Counts

	Model 5 _{MLE}		Model 6 _{2step}	
	Selection	Outcome	Selection	Outcome
Civil War Intensity	−0.24 (0.20)		−0.11 (0.21)	
Forced Migration _{t-1}	0.50** (0.15)		0.35* (0.15)	
Interstate War	0.19 (0.29)		0.23 (0.32)	
Conflict Duration	−0.02 (0.01)		−0.02 (0.02)	
log(Intercommunal Deaths)	0.00 (0.03)		0.01 (0.04)	
log(Natural Disaster Victims)	0.03 (0.04)		−0.04 (0.03)	
log(Civilian Deaths)	0.11** (0.03)		0.11** (0.03)	
Ethnic War		0.17** (0.06)		0.17** (0.06)
Population Density		0.00 (0.00)		0.00 (0.00)
Island		−0.34** (0.13)		−0.39** (0.13)
Civil War _{Neighborhood}		0.05 (0.07)		0.05 (0.07)
Intcom.Viol _{Neighborhood}		−0.07 (0.06)		−0.05 (0.07)
Civil.Targ _{Neighborhood}		−0.04 (0.08)		−0.04 (0.08)
Borders		−0.01 (0.01)		−0.02 (0.01)
Ref:Forced Migrants _{t-1}		0.25** (0.05)		0.30** (0.08)
σ		0.41**		
ρ		0.60**		
invMillsRatio				0.50**
Adj. R ²		0.14		0.13
Num. obs.	227	365	227	365

Note:

*p<0.05; **p<0.01

TABLE C.3: Conditional Effects Models with Log Counts

	Model 7 _{MLE}		Model 8 _{2step}	
	Selection	Outcome	Selection	Outcome
Civil War Intensity	−0.19 (0.21)		−0.11 (0.21)	
Forced Migration _{t-1}	0.45** (0.15)		0.35* (0.15)	
Interstate War	0.21 (0.31)		0.23 (0.32)	
Conflict Duration	−0.02 (0.02)		−0.02 (0.02)	
log(Intercommunal Deaths)	0.00 (0.04)		0.01 (0.04)	
log(Natural Disaster Victims)	−0.03 (0.04)		−0.04 (0.03)	
log(Civilian Deaths)	0.10** (0.03)	−0.05 (0.02)	0.11** (0.03)	−0.05 (0.02)
Civilian Deaths X Ethnic War		0.18* (0.08)		0.19* (0.07)
Ethnic War		−0.03 (0.10)		0.17** (0.06)
Population Density		0.00 (0.00)		0.00 (0.00)
Island		−0.41** (0.13)		−0.41** (0.13)
Civil War _{Neighborhood}		0.05 (0.07)		0.05 (0.07)
Intcom.Viol _{Neighborhood}		−0.04 (0.06)		−0.05 (0.06)
Civil.Targ _{Neighborhood}		−0.06 (0.08)		−0.04 (0.08)
Borders		−0.02 (0.01)		−0.02 (0.01)
Ref:Forced Migrants _{t-1}		0.04** (0.01)		0.30** (0.08)
σ		0.37**		
ρ		0.40*		
invMillsRatio				0.17*
Adj. R ²		0.14		0.15
Num. obs.	227	365	227	365

Note:

*p<0.05; **p<0.01

TABLE C.4: Three-Stage (1) Heckit

	Selection _{Stage 1}
(Intercept)	−0.75 (0.70)
log(Population)	0.16* (0.07)
Polity	−0.00 (0.02)
Polity ²	0.00 (0.00)
Regime Stability	−0.20 (0.23)
Ethnic Fractionalization	0.38 (0.37)
log(Mountainous Terrain)	0.03 (0.07)
Infant Mortality	0.33* (0.14)
Civil War _{Neighborhood}	−0.13 (0.08)
Peaceyears	−2.57** (0.24)
Peaceyears2	0.57** (0.07)
Peaceyears3	−0.04** (0.01)
Slow Growth	0.18 (0.19)
AIC	266.86
Log Likelihood	−120.43
Num. obs.	2,720

** $p < 0.01$, * $p < 0.05$

TABLE C.5: Three-Stage (2) Heckit

	Selection _{Stage 2}	Outcome
Civil War Intensity	−0.10 (0.20)	
Forced Migration _{t-1}	0.41 (0.15)	
Interstate War	0.15 (0.30)	
Conflict Duration	−0.03 (0.02)	
log(Intercommunal Deaths)	0.05 (0.18)	
log(Natural Disaster Victims)	−0.09 (0.15)	
log(Civilian Deaths)	0.70** (0.03)	
Ethnic War		0.17** (0.06)
Population Density		0.00 (0.00)
Island		−0.40 (0.13)
Civil War _{Neighborhood}		0.03 (0.07)
Intcom.Viol _{Neighborhood}		−0.04 (0.06)
Civil.Targ _{Neighborhood}		−0.05 (0.07)
Borders		−0.02 (0.01)
Ref:Forced Migrants _{t-1}		0.30** (0.06)
invMillsRatio	−0.03	0.56**
Adj. R ²		0.14
Num. obs.	365	227

Note: *p<0.05; **p<0.01

Bibliography

- Adelman, Howard. 1998. "Why refugee warriors are threats." *Journal of Conflict Studies* 18(1).
- Bartusevičius, Henrikas. 2016. "Introducing the Categorically Disaggregated Conflict (CDC) dataset." *Conflict Management and Peace Science* 33(1):89–110.
- Bhavnani, Ravi and Dan Miodownik. 2009. "Ethnic polarization, Ethnic Salience, and Civil War." *Journal of Conflict Resolution* 53(1):30–49.
- Bohnet, Heidrun. 2012. Large Refugee Populations, Resource Scarcity and Conflict. In *Annual Conference of the European Political Science Association*. p. 21–23.
- Bohnet, Heidrun, Fabien Cottier and Simon Hug. 2013. "Conflict-induced IDPs and the Spread of Conflict." *Presentation at the European Political Science Association* .
- Boserup, Ester. 1990. "Economic and Demographic Relationships in Development." *Population and Development Review* 16(4):775–779.
- Buhaug, Halvard. 2010. "Dude, where's my conflict? LSG, relative strength, and the location of civil war." *Conflict Management and Peace Science* .
- Carter, David B and Curtis S Signorino. 2010. "Back to the future: Modeling time dependence in binary data." *Political Analysis* 18(3):271–292.
- Caselli, Francesco and Wilbur John Coleman. 2013. "On the theory of ethnic conflict." *Journal of the European Economic Association* 11(s1):161–192.
- Cederman, Lars-Erik, Andreas Wimmer and Brian Min. 2010. "Why Do Ethnic Groups Rebel? New Data and Analysis." *World Politics* 62(01):87–119.
- Chi, Guangqing and Jun Zhu. 2008. "Spatial Regression Models for Demographic Analysis." *Population Research and Policy Review* 27(1):17–42.
- Davenport, Christina, Will Moore and Steven Poe. 2003. "Sometimes you just have to leave: Domestic threats and forced migration, 1964-1989." *International Interactions* 29(1):27–55.

- Downes, Alexander B. 2006. "Desperate Times, Desperate Measures: The Causes of Civilian Victimization in War." *International Security* 30(4):152–195.
- Drabo, Alassane and Linguère Mbaye. 2011. "Climate change, Natural disasters and Migration: An Empirical Analysis in Developing Countries."
- Eck, Kristine. 2009. "From Armed Conflict to War: Ethnic Mobilization and Conflict Intensification*." *International Studies Quarterly* 53(2):369–388.
- Eck, Kristine and Lisa Hultman. 2007. "One-Sided Violence Against Civilians in War Insights from New Fatality Data." *Journal of Peace Research* 44(2):233–246.
- Ehrlich, Paul, R. 1968. "The Population Bomb." *Ballantine, New York*.
- Engelhardt, Michael J. 1992. "Democracies, Dictatorships and Counterinsurgency: Does Regime Type Really Matter?" *Journal of Conflict Studies* 12(3).
- Esteban, Joan and Debraj Ray. 2008. "Polarization, fractionalization and conflict." *Journal of peace Research* 45(2):163–182.
- Fearon, James D. 1995. "Rationalist explanations for war." *International organization* 49(03):379–414.
- Fearon, James D. 1998. "Commitment Problems and the Spread of Ethnic Conflict. The International Spread of Ethnic Conflict." *Princeton, NJ, Princeton University Press* 52:269–305.
- Fearon, James D. 2004. "Ethnic Mobilization and Ethnic Violence." *Barry R. Weingast and*.
- Fearon, James D and David D Laitin. 2003. "Ethnicity, Insurgency, and Civil war." *American Political Science Review* 97(01):75–90.
- Ferris, Elizabeth G. 1993. *Beyond Borders: Refugees, Migrants and Human Rights in the Post-Cold War Era*. WCC publications Geneva.
- Fisk, Kerstin. 2014. "Refugee Geography and the Diffusion of Armed Conflict in Africa." *Civil Wars* 16(3):255–275.
- Fjelde, Hanne and Lisa Hultman. 2010. Weakening Your Enemy: Constituencies and the Location of Violence against Civilians in Africa, 1989-2006. In *SGIR 7th Pan-European International Relations Conference*.
- Forsberg, Erika. 2009. "Refugees and Intrastate Armed Conflict: A Contagion Process Approach."

- Franzese, Robert J, Jude C Hays and Lena M Schaffer. 2010. Spatial, Temporal, and Spatiotemporal Autoregressive Probit Models of Binary Outcomes: Estimation, Interpretation, and Presentation. In *Annual Conference of the American Political Science Association*.
- Garcia-Montalvo, Jose and Marta Reynal-Querol. 2004. "Ethnic Polarization, Potential Conflict, and Wivil Wars." *Potential Conflict, and Civil Wars* .
- Gleditsch, Nils Petter et al. 2002. "Armed conflict 1946-2001: A new dataset." *Journal of Peace Research* 39(5):615–637.
- Hall, Bronwyn H. 2002. "Notes on Sample Selection Models." .
URL: <http://eml.berkeley.edu//bhhall/e244/sampsel.pdf>
- Harff, Barbara. 2003. "No Lessons Learned from the Holocaust? Assessing Risks of Genocide and Political Mass Murder Since 1955." *American Political Science Review* 97(01):57–73.
- Hassner, Ron Eduard. 2009. *War on Sacred Grounds*. Cornell University Press.
- Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47(1):153–161.
- Hegre, Håvard, Gudrun Østby and Clionadh Raleigh. 2009. "Poverty and Civil War Events: A Disaggregated Study of Liberia." *Journal of Conflict Resolution* 53(4):598–623.
- Homer-Dixon, Thomas and Jessica Blitt. 1998. *Ecoviolence: Links among Environment, Population, and Security*. Rowman & Littlefield Publishers.
- Kalyvas, Stathis. 2001. "New and old civil wars." *World politics* 54(1):99–118.
- Kalyvas, Stathis N. 2005. "Warfare in civil wars." *Rethinking the Nature of War* pp. 88–108.
- Kaufmann, Chaim. 1996. "Possible and Impossible Solutions to Ethnic Civil Wars." *International Security* 20(4):136–175.
- Lischer, Sarah Kenyon. 2006. *Dangerous Sanctuaries: Refugee Camps, Civil war, and the Dilemmas of Humanitarian Aid*. Cornell University Press.
- Loescher, Gil and James Milner. 2005. "Security Implications of Protracted Refugee Situations." *The Adelphi Papers* 45(375):23–34.
- Malthus, Thomas Robert. 1798. *An Essay on the Principle of Population, as it affects the Future Improvement of Society: with remarks on the Speculations of Mr. Godwin, M. Condorset, and other writers*. J. Johnson.
- Marfleet, Philip. 2007. "Refugees and history: why we must address the past." *Refugee Survey Quarterly* 26(3):136–148.

- Marshall, Monty G. and Robert T. Gurr. 2014. "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2013." .
URL: <http://www.systemicpeace.org/polity/polity4.htm>
- Maxwell, John W and Rafael Reuveny. 2000. "Resource Scarcity and Conflict in Developing Countries." *Journal of Peace Research* 37(3):301–322.
- Moore, Will H and Stephen M Shellman. 2004. "Fear of Persecution Forced Migration, 1952-1995." *Journal of Conflict Resolution* 48(5):723–745.
- Moore, Will H and Stephen M Shellman. 2007. "Whither Will They Go? A Global Study of Refugees' Destinations, 1965-1995." *International Studies Quarterly* 51(4):811–834.
- Neumayer, Eric. 2005. "Asylum Recognition Rates in Western Europe Their Determinants, Variation, and Lack of Convergence." *Journal of Conflict Resolution* 49(1):43–66.
- Osborn, Fairfield. 1948. "Our Plundered Planet." .
- Ousey, Graham C, Pamela Wilcox and Bonnie S Fisher. 2011. "Something old, Something New: Revisiting Competing Hypotheses of the Victimization-offending Relationship among Adolescents." *Journal of Quantitative Criminology* 27(1):53–84.
- Phillipson, Coleman. 1911. *The international law and custom of ancient Greece and Rome*. Vol. 1 Macmillan and Company, limited.
- Posen, Barry R. 1993. "The Security Dilemma and Ethnic Conflict." *Survival* 35(1):27–47.
- Rabben, Linda. 2011. *Give refuge to the stranger: the past, present, and future of sanctuary*. Left Coast Press.
- Raleigh, Clionadh and Henrik Urdal. 2007. "Climate Change, Environmental Degradation and Armed Conflict." *Political Geography* 26(6):674–694.
- Reyna, Stephen P. 2010. "The Disasters of War in Darfur, 1950–2004." *Third World Quarterly* 31(8):1297–1320.
- Reynal-Querol, Marta. 2002. "Ethnicity, Political Systems, and Civil Wars." *Journal of Conflict Resolution* 46(1):29–54.
- Rüegger, Seraina and Heidrun Bohnet. 2015. "The Ethnicity of Refugees (ER): A new dataset for understanding flight patterns." *Conflict Management and Peace Science* .
- Salehyan, Idean. 2008. "The externalities of civil strife: Refugees as a source of international conflict." *American Journal of Political Science* 52(4):787–801.

- Salehyan, Idean and Kristian Skrede Gleditsch. 2006. "Refugees and the spread of civil war." *International Organization* 60(02):335–366.
- Sambanis, Nicholas. 2001. "Do Ethnic and Nonethnic Civil Wars Have the Same Causes? A Theoretical and Empirical Inquiry (Part 1)." *Journal of Conflict Resolution* 45(3):259–282.
- Sarkees, Meredith Reid and Frank Wayman. 2010. "Resort to War: 1816 - 2007." *Conflict Management and Peace Science* .
- Schechtman, Joseph B. 1964. *The Refugees in the World: Displacement and Integration*. Barnes.
- Schmeidl, Susanne. 2001. "Conflict and forced migration: a quantitative review, 1964-1995." *Global migrants, global refugees: Problems and solutions* pp. 62–94.
- Shaver, Andrew and Yang-Yang Zhou. 2015. "Questioning Refugee Camps as Sources of Conflict." .
- Sundberg, Ralph. 2009. Revisiting One-sided Violence—A Global and Regional Analysis. In *States in Armed Conflict*, ed. Lotta Harbom and Ralph Sundberg. Department of Peace and Conflict Research.
- Sundberg, Ralph, Kristine Eck and Joakim Kreutz. 2012. "Introducing the UCDP Non-state Conflict Dataset." *Journal of Peace Research* 49(2):351–362.
- Theisen, Ole Magnus. 2008. "Blood and Soil? Resource Scarcity and Internal Armed Conflict Revisited." *Journal of Peace Research* 45(6):801–818.
- Thompson, W.S. 1929. "Population." *American Journal of Sociology*. (34).
- Tilly, Charles. 1975. "The Formation of National States in Western Europe." *Princeton University Press* .
- Urdal, Henrik. 2005. "People vs. Malthus: Population Pressure, Environmental Degradation, and Armed Conflict Revisited." *Journal of Peace Research* 42(4):417–434.
- Valentino, Benjamin, Paul Huth and Dylan Balch-Lindsay. 2004. "'Draining the sea': Mass killing and guerrilla warfare." *International Organization* 58(02):375–407.
- Vernant, Jacques. 1953. *The Refugee in the Post-War World*. Yale University Press.
- Vogt, Manuel et al. 2015. "Integrating Data on Ethnicity, Geography, and Conflict The Ethnic Power Relations Data Set Family." *Journal of Conflict Resolution* .
- Vogt, William. 1948. *Road to Survival*. New York: William Sloane Associates, Inc.

- Weidmann, Nils B and Michael D Ward. 2010. "Predicting Conflict in Space and Time." *Journal of Conflict Resolution* 54(6):883–901.
- Wimmer, Andreas, Lars-Erik Cederman and Brian Min. 2009. "Ethnic politics and armed conflict: a configurational analysis of a new global data set." *American Sociological Review* 74(2):316–337.
- Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT press.
- Zolberg, Aristide R, Astri Suhrke and Sergio Aguayo. 1992. *Escape from violence: Conflict and the refugee crisis in the developing world*. Oxford University Press on Demand.