Refugee Cascades and the Spatial Spillover of Fleeing in Civil War

Benjamin Laughlin

December 31, 2016

Abstract

I conceptualize refugee movements as the result of decisions by sophisticated civilians in response to new information about the risk of victimization in war. One major source of information available to civilians living in conflict zones is observing refugees fleeing. This can result in an information cascade, causing a refugee cascade, in which waves of refugees fleeing cause other civilians to increase their beliefs about the risk, increasing the numbers of subsequent refugees. I construct a geocoded village-day level dataset of refugee flows, violence against civilians, and the actions of armed groups during the Kosovo war. I develop an instrumental variables estimation strategy using the spatial network of villages connected by roads and the fact that refugees fled toward a single border crossing to estimate the causal spillover effect of refugees fleeing. I find that on average a refugee fleeing causes one additional civilian to flee, which is larger than the effect of violence, the dominant explanation for refugee movements in the literature.

1 Introduction

Why do civilians living in war zones choose to leave and become refugees? The answer may seem obvious, but in most wars, only a small minority of civilians ever flee as refugees. Even after four years of war in Syria, widely publicized for the refugee crisis it has caused, only about 20 percent of the population has left the country as refugees.[1] Even in conflicts noted for their brutality against civilians and ethnic cleansing, like Bosnia and Kosovo, only a minority ever left. Going back to Hirschmann (1970), scholars have considered the option of "exit" available to citizens who dislike the quality or policies of their government. Civil wars represent arguably the most extreme form of poor quality governance, yet there is not always a quick rush to exit.

This paper investigates how civilians acquire information, learn about their risk of becoming a victim of violence, and use that to make a decision about fleeing or staying in the conflict zone. We know little about how civilians living in conflict zones acquire information about their risk, in spite of the recognized importance of information in civil war environments (Kalyvas, 2006). "Mistakes" in this decision—fleeing when it was safe to stay or staying when it was too dangerous—have enormous consequences. Remaining when it was more dangerous than anticipated leads to excess civilian casualties. Large numbers of refugees fleeing when they would have been safe at home not only causes an enormous human cost on the refugees themselves, who lose their assets and often end up in poor conditions or subjected to discrimination, but may have large economic[2] and social costs on the countries refugees flee to (Dancygier and Laitin, 2014; Salehyan and Gleditsch, 2006; Salehyan, 2008).

A civilian living in a war zone has a difficult problem. She is forced to trade off the risk to her personal security with the likely loss of personal property associated with fleeing (Davenport,

[1] In February 2016 the UN estimated that there were 4.6 million refugees from Syria (http://www.unocha.org/syria) out of a pre-war population of 22.9 million.

[2] Germany, for example, estimates that it costs 12,500 euros per refugee per year (http://www.dw.com/en/what-helping-refugees-costs-germany/a-18693996).

Moore and Poe, 2003; Edwards, 2009; Melander and Öberg, 2006; Moore and Shellman, 2004, 2006, 2007; Schmeidl, 1997). The personal costs to fleeing one's home can be large. Ibáñez and Vélez (2008) estimate that the welfare loss of internal displacement in Columbia is 37% of the net present value of rural lifetime aggregate consumption. Becoming a refugee in another country, where language and cultural barriers may present additional difficulties, is potentially even more costly.

Civil wars are environments where information is scarce and civilians are locally informed (Walter, 2009; Kalyvas, 1999). In such an information-poor environment a civilian needs to use what little sources of information available to her to make a potentially life and death decision. Sources of information available to civilians are friends and acquaintances who can pass on what they know or have heard and observation of the behavior of others nearby.

Hearing information or rumors from other civilians nearby or even just seeing them or evidence of their passage provides a civilian with some information. In the absence of better information, the actions of others can have a substantial impact on an individual's own perception of the risk of violence and therefore on her decision to flee and become a refugee. If this civilian flees, she can further impact the information and decision making processes of other civilians, creating a refugee cascade caused by the information transmitted to other civilians by a stream of refugees fleeing.

This paper shows that refugees' fleeing causes additional fleeing from nearby places, providing evidence that civilians learn from their neighbors' decisions about fleeing, which in turn are affected by their neighbors. However, most studies of the determinants of civilian displacement in civil war have focused only on the impact of local violence and economic factors on the decision to flee. Those that have considered information in civilian's decision making processes have only focused on information about the journey and destination learned from refugees who have already fled (e.g., Davenport, Moore and Poe, 2003; Edwards, 2009; Moore and Shellman, 2004, 2007). As a result, our understanding of displacement in civil war remains incomplete.

This study finds that refugees fleeing has substantial local spillover effects. Using a fine-grained, geocoded panel dataset on the Kosovo war compiled from a number of different sources and a spatial instrumental variables research design to find the causal effect of refugees fleeing from neighboring villages on fleeing. The spatial nature of the data and the direction of the migration patterns allows violence in the neighboring villages of neighboring villages that are not also your neighbor to act as instrumental variables for the spillover effect of fleeing in some geographic directions, but not others. Specifically, it uses the fact that refugees had to travel to a single border post on the Albanian border. Using the direction of their travel from their home villages, we know which other villages refugees could potentially have passed near and conveyed information on the risk of violence.

Although the analysis in this paper is based on one conflict, the results can be applied more broadly. When refugees fleeing can affect other civilians' beliefs about risk and decision making, the spatial and temporal distribution of violence matters. The same amount of violence in different places will result in different numbers of refugees because cascades of refugees generated by information spillovers will differ.

This paper contains two main contributions. First, it contributes to the literature on the determinants of displacement in civil war, where I emphasize the need to consider the dynamics of the conflict and information flows on civilians' decisions to flee and become refugees and make a theoretical contribution, arguing that cascades of refugees occur in civil wars and that changes in the spatial distribution of violence can radically alter the number of refugees that result from the conflict. Second, it contributes to the methods literature by causally identifying spatial spillover effects. Until now, the quantitive methods literatures on spatial models and causal inference have remained largely separate.

2 Displacement in Civil War

A large literature uses cross-national data to investigate the determinants of refugee flows (see, e.g., Davenport, Moore and Poe, 2003; Melander and Öberg, 2006, 2007; Moore and Shellman, 2004, 2006; Schmeidl, 1997). The most robust finding of this literature is that violence against civilians is strongly associated with refugee flows (Apodaca, 1998; Davenport, Moore and Poe, 2003; Gibney, Apodaca and McCann, 1996; Jonassohn, 1993; Moore and Shellman, 2004, 2006; Rummel, 1994; Schmeidl, 1997; Weiner, 1998). However, analysis within a single country has several advantages over cross-national studies. It allows the evaluation of hypotheses that cross-national studies cannot easily consider—like the effect of variations in violence or information within a war—and this allows the consideration of more complex theories about linkages between regions within a country (Snyder, 2001). By increasing the number of observations within a case, a within-country study also allows for easier casual inference because difficult to measure country-level variables like history and culture are held constant (King, Keohane and Verba, 1994).

More recently, several within-country studies of the determinants of civilian displacement in civil wars have been conducted using surveys of internally displaced people in Nepal and Colombia (Adhikari, 2013, 2012; Bohra-Mishra and Massey, 2011; Engel and Ibáñez, 2007; Lozano-Gracia et al., 2010; Ibáñez and Vélez, 2008; Steele, 2011). These generally conclude that local violence and economic factors are associated with fleeing from conflict zones. Several of them find that increases in income reduces the likelihood of fleeing for the same violence-level. Balcells and Steele (2012) use within-country data constructed at small geographical units in Spain and Colombia. However, they do not consider civilians living in conflict zones to have agency, instead theorizing as to why armed groups would unilaterally choose to displace civilians that have opposing political viewpoints. The mechanism described by Balcells and Steele (2012), considering two non-ethnic conflicts, cannot explain the variation in refugee flows from the Kosovo War. In Kosovo, virtually the entire ethnic Albanian majority was opposed to the openly hostile Milošević government after the abolition of Kosovo's autonomy in 1989[3] and these feelings has only become stronger after a decade of repression. Therefore, the uniformly opposed political views of ethnic Albanians in Kosovo cannot explain why many ethnic Albanians civilians fled Kosovo during the war, but most did not.

Factors other than local violence and the economy may be important for migration decisions. Social networks and the communication across these ties may play a role in fleeing. Individuals are embedded in preexisting social networks in their communities (Adhikari, 2013). These ties may make them more likely to stay with the people they know, but they also allow the transmission of information. Besides preexisting social networks, refugees meet and talk to other civilians while they are fleeing (Edwards, 2009). Spitzer (2015) shows that nearby pogroms in the Russian Empire were not the main driver of late-nineteenth and early-twentieth century Jewish immigration to the United States and argues that information transmitted through "chain-migration networks" is a bigger cause. That is, social networks allowed potential migrants to gain information that increased their likelihood of choosing to migrate themselves.

3 Theory

Civilians living in a war zone face a difficult problem. There is a trade-off in the risk of violence and the likely loss of personal property associated with fleeing (Davenport, Moore and Poe, 2003; Edwards, 2009; Melander and Öberg, 2006; Moore and Shellman, 2004, 2006, 2007; Schmeidl, 1997). As Adhikari (2013) notes, "people make a decision to flee or stay even under highly dangerous circumstances." While the primary concern of most civilians during civil war is to

[3]For example, when Albanian nationalist leaders asked ethnic Albanians to boycott the 1991 census in protest of Serb dominance, virtually all did (Bieber, 2015).

remain alive, people often accept risks in return for other things of value to them.[4]

In the case of fleeing and becoming a refugee, the personal costs can be enormous. Many refugees can expect to never return to their former homes. Ibáñez and Vélez (2008) estimate that the welfare loss of internal displacement in Columbia is 37% of the net present value of rural lifetime aggregate consumption. Becoming a refugee in another country is potentially even more costly, where language and cultural barriers may be difficult. Wiesner (1988) points out that during the Vietnam War, the amount civilians would lose if they fled their villages determined how long they remained. Land-owning peasants were the least likely to leave because they had nothing else.

Nordstrom (1997) quotes a Mozambican man who illustrated the dilemma facing civilians living in war zones: "if you try to protect yourself and flee you lose everything. If you stay you may keep your possessions and lose your life." Because of the high costs to becoming a refugee, not everyone flees at the first sign of trouble. Instead, individuals choose to flee only when their beliefs about the risk of becoming victims of violence are high enough to outweigh the loss of their material assets or the personal cost of fleeing.

In order to make the difficult decision to flee your home toward an largely unknown foreign country, a civilian living in a conflict zone needs to assess her personal risk of becoming a victim of violence. Unfortunately civil wars are environments where information is scarce and civilians are locally informed (Walter, 2009; Kalyvas, 1999). It is difficult to know what is happening in another place, even if it is not actually very far away. As wars destroy telecommunications and transportation infrastructure and make even short journeys potentially dangerous, outsiders living abroad often have a much better idea of the state of the conflict than do those living within it. In such an information-poor environment a civilian needs to use what little sources of information available to her to make a potentially life and death decision. Local social networks—friends and acquaintances who can pass on what they know or have heard and observation of the behavior of others nearby—provide opportunities to gain new information and better assess the risk of remaining.[5]

Observations that individuals fled at a certain time implies that their belief about the probability of becoming victims of violence had risen sufficiently high to change their mind about remaining in their homes. New information is needed for individuals to change their beliefs about the risk of violence. People can acquire information either from combatants or from other civilians. They learn from combatants directly by observing violence against their neighbors. Civilians from other villages who are themselves fleeing past these individuals' villages can tell them about violence that they witnessed or heard about, changing individuals' beliefs about their own risk. They can also observe other civilians fleeing and update their beliefs about their personal risk of violence that way. The lack of good information during a war means that any information individuals do acquire has a larger impact on their decisions than if they were in a situation with more access to information about the risk they face.

One major source of information available to civilians is from other civilians who are fleeing from their own villages. This information can be conveyed by talking to these refugees as they pass and learning about the violent events that may have motivated their flight or just by seeing them pass nearby. Even without actually communicating with the refugees, knowing that others are fleeing is enough to increase an individual civilian's belief about the probability of being a victim of violence and therefore increase her likelihood of fleeing herself. This means that every subsequent refugee who passes by increases the expected number of civilians who will flee from their homes. If this is enough to motivate them to flee themselves, these new civilians fleeing

[4]It is unclear how different levels of risk aversion might affect the likelihood of fleeing a conflict zone and becoming a refugee. The uncertainty involved in leaving for an unknown place might deter risk averse individuals from migrating (Fischer, Martin and Straubhaar, 1997). However, sufficiently high expectations of the risk of violence may cause risk averse individuals to become refugees in spite of this.

[5]Acquiring information about the journey and destination from other civilians in order to make decisions about fleeing or staying has been considered in the literature, but acquiring information from other civilians within the conflict about the risk of remaining has not been discussed (Davenport, Moore and Poe, 2003; Edwards, 2009; Moore and Shellman, 2004, 2007).

then can influence the fleeing decisions of civilians in subsequent villages.

This situation, where observing someone fleeing can cause an individual to flee herself, which can cause someone else to subsequently flee, is an information cascade. Banerjee (1992) and Bikhchandani, Hirshleifer and Welch (1992, pp. 992) started a theoretical literature in economics on social learning, in which individuals learn by observing the behavior of others. Calling this phenomenon an information cascade or herding, Bikhchandani, Hirshleifer and Welch (1992, 992) define it as occurring "when it is optimal for an individual, having observed the actions of those ahead of him, to follow the behavior of the preceding individual without regard to his own information." This theoretical literature concludes that such situations are rational, but result in inefficient outcomes (Bikhchandani, Hirshleifer and Welch, 1998). One important result is that cascades can be based on very little information. Once a cascade begins, individuals ignore their own information in favor of the information from observing others fleeing. This means that the right initial event can result in large numbers of refugees leaving their homes even though the actual probability of most of these individuals imminently experiencing violence has not changed. The actual risk to an individual is local, but information cascades can propagate information far beyond its actual useful range.

4 Case: The Kosovo War

The war in Kosovo in 1999 provides an ideal case to investigate how information about risk is transmitted and how the affects civilians' decisions to flee from the conflict zone. This war resulted in over 600,000 refugees out of a population of 2 million[6] and over 40,000 deaths. The war stated as a low-level insurgency by the Albanian-nationalist Kosovo Liberation Army (KLA), but rapidly expanded in 1998 as the Yugoslav army heavily repressed the ethnic Albanian population of Kosovo. Concerned about ethnic cleansing and large numbers of ethnic Albanian refugees from Kosovo, NATO initiated a bombing campaign that lasted from March 24, 1999 to June 10, 1999, when the war ended. While ethnic Albanians make up the majority of the victims of the war, there was also violence against Serb civilians by the KLA and fleeing from Kosovo by Serb civilians. I restrict this analysis to fleeing by and violence against ethnic Albanian civilians. By considering only fleeing by one ethnic group, some factors identified as potential causes of displacement, such as which armed faction the civilian supports, are held constant.

One factor that makes the case of Kosovo ideal is that the war happened prior to the introduction of cell phones, so when landlines were cut off early in the war, communication became local. The widespread use of cell phones, like is happening in the Syrian Civil War, would make identifying social networks much more difficult because they could no longer be based on spatial proximity. During the Kosovo war, communicating with people who were not physically nearby was impossible. Another nice feature of the Kosovo case is that all refugees in the data were fleeing in the same direction: toward a single border crossing with Albania. However, there was little contact with Albania by these civilians prior to becoming refugees because international travel was difficult in Yugoslavia.

Comprehensively measuring fleeing from war zones is difficult due to the chaos of the conflict situation and the often surreptitious nature of the fleeing. This is likely a reason for the dearth of quantitative research on civilians fleeing within conflicts. Most conflicts create not just refugees, civilians who flee to other countries, but also internally displaced people, who leave their homes, but remain in their own country. The internally displaced are often harder to count, interview, or survey, but they are of equal theoretical interest to questions about why civilians flee their homes. The Internal Displacement Monitoring Center estimates nearly twice as many internally displaced people as refugees in the world in 2014, creating a problem for research using only data on refugees.[7] Kosovo provides an ideal case because of its small size and because it had

[6] https://migration.ucdavis.edu/mn/more.php?id=1801

[7] For 2014 they estimate 38 million IDPs and 19.5 million refugees. See http://www.internal-displacement.org/global-figures.

Table 1: Descriptive Statistics

	Level	Obs.	Mean	Std. Dev.	Min.	Max.
Refugees	village-day	89,404	1.9	77.5	0	8874
Civilians killed	village-day	89,404	0.06	1.49	0	274
KLA killed	village-day	89,404	0.01	0.25	0	28
Army killed	municipality-day	1798	0.11	0.53	0	6
Battles	municipality-day	1798	0.04	0.51	0	20
NATO airstikes	municipality-day	1798	0.16	0.38	0	2
Population	village	1442	1098.0	4307.4	5	108,083
Albanians	village	1442	850.4	3057.5	0	75,803
Serbs	village	1442	145.3	660.0	0	16,898
Distance to border (m)	village	1442	97,470	37,239.3	1,215	175,600
Distance to highway (m)	village	1442	1786	1593.8	0.62	10,110

Table 2: Number of Neighboring Villages

	Neighborhood distance				
	5 km	10 km	15 km	20 km	25 km
Neighbors (all directions)	11.82	44.30	92.88	154.39	226.57
Neighbors (toward border)	7.54	33.25	75.32	130.9	197.34

a sympathetic neighbor in Albania, willing to accept unlimited numbers of coethnic refugees.[8] These factors allowed most people who fled their homes to exit the country. Even though I, like others, do not count internally displaced people, this is much less of a problem in this conflict than in most others.

Kosovo is geographically small and surveys conducted in refugee camps show that most refugees left the country the same day they fled their homes (Iacopino et al., 2001; Ball, 2000).

5 Data

I combine data from multiple sources to construct a village-day panel dataset of refugees fleeing, violence against civilians, and violence between combatants between March 28, 1999 and May 28, 1999, the period that saw the greatest number of refugees flee from Kosovo. Records of refugees fleeing recorded on Albanian border records, violence against civilians collected by human rights organizations, and information on the actions of armed groups introduced as evidence by the International Criminal Tribunal for the former Yugoslavia (ICTY) Office of the Prosecutor, and demographic data from the census were combined based on named locations and dates. Descriptive statistics of the variables used in the analysis are shown in table 1. Every village was geocoded so that neighboring villages could be determined. The number of neighboring villages for different neighborhood threshold distances is shown in table 2.

[8]The Yugoslav army did briefly close access to the Albanian border several times during the war and reports by refugees of being forced to bribe Yugoslav soldiers to be allowed to reach the border are common.

5.1 Refugees fleeing

I measure refugees fleeing during the Kosovo war using border crossing records collected at the Morina border crossing point with Albania. Records collected by Albanian border guards at the Morina border crossing between Kosovo and Albania were photographed by Ball et al. (2002), a consultant for the ITCY prosecutor. This was the only border crossing between Kosovo and Albania open during the war, the primary recipient country of refugees during the war. Albanian officials at the border recorded the size of the party crossing the border, the date, and the origin village or town of the group. Each of the 19,126 records represents a single individual, household or group that crossed the border. These groups range in size from 1 to 1961 people, with a mean of 14.45. The Yugoslav army heavily patrolled the border with Albania to prevent KLA incursions, so few refugees crossed into Albania using other routes. Lines of people waiting to cross often extending back several kilometers from the border.[9] As Klingner and Silva (2013, pp. 157) notes, "these border crossing records are remarkably complete. Nearly all of the refugee flow from Kosovo into Albania went through this one small border post, and Albanian officials there attempted to document each crossing." Figure 1 shows the geographic distribution of the origins of refugees.

5.2 Violence

Data on the killing and disappearances of civilians during the war are from the "Kosovo Memory Book, 1998-2000," a comprehensive list of every known death or disappearance in Kosovo due to the war that was compiled by the Humanitarian Law Centre, a Belgrade and Prishtina based human rights organization. It lists 13,549 people by name who were killed or are missing from the conflict as well as the date, location, and known details of the incident. For this analysis I restrict my measure of violence to ethnic Albanian civilians killed or known to have disappeared between March and June 1999, to correspond to the data on ethnic Albanian refugees. I dropped any reported violence for which village-level location information was not available. This included violence with only a municipality-level location identified and that which had an even more vague location (e.g., "in the mountains"). This affected 4.6 percent of the records. I also dropped 204 reported deaths for which only the year is known. Figure 2 shows the distribution of violence against civilians throughout Kosovo. All areas of Kosovo encountered killings of civilians by Yugoslav forces.

5.3 Combat

Data on NATO airstrikes and KLA activity are only available at the municipality-level. There were 29 municipalities in Kosovo in 1999 and this data was matched to all villages within the geographic boundaries of the municipality. Data on the 364 NATO airstrikes are from Arkin (2000). Data on both battles and Yugoslavian casualties due to KLA activity are from Ball et al. (2002), based on data originally from the ICTY Office of the Prosecutor. This data omits several incidents included in the original list of KLA attacks because of ITCY confidentiality rules.

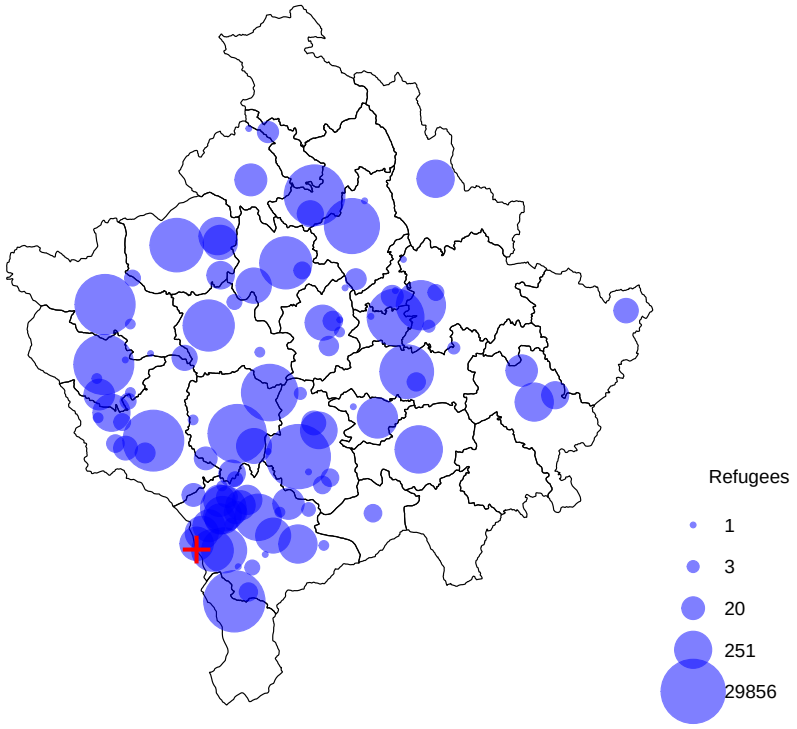
5.4 Demographics

I use the 1981 census of Yugoslavia for locality-level demographic data, with the 1445 villages, towns, or cities in this census serving as the units of analysis.[10] The 1991 census was not used because of a large-scale boycott of the census by ethnic Albanians that was promoted by Albanian nationalist groups in Kosovo. Every village was geocoded so that neighboring villages could be determined.

[9] http://reliefweb.int/report/albania/more-kosovar-refugees-pour-albania

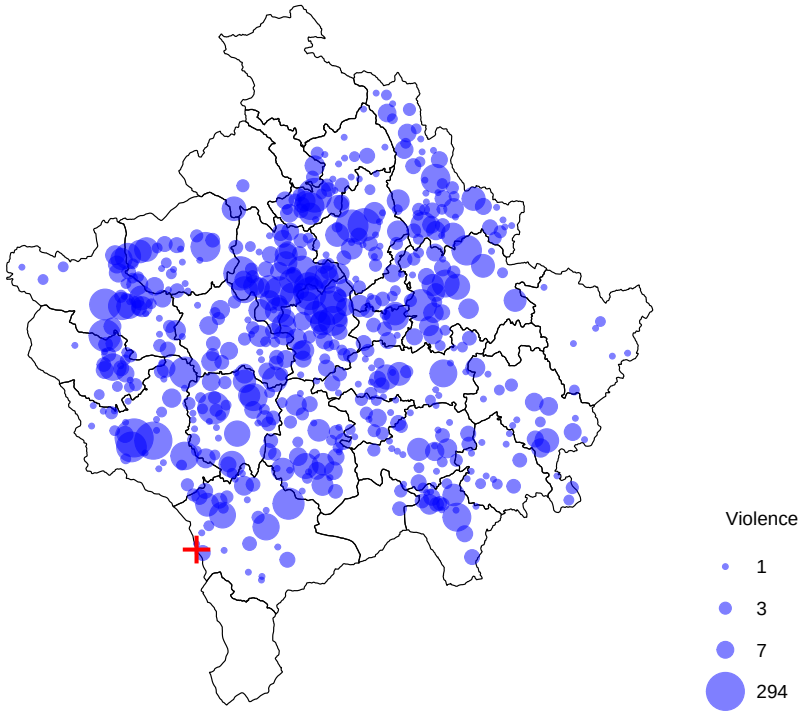
[10] In three cases neighboring villages with the same name prefixed by "Upper" and "Lower" were combined because it was impossible to determine which village border records or reports of violence were referring to.

Figure 1: Origins and Magnitudes of Refugee Flows from Kosovo



The origins of refugees who entered Albania between March 28, 1999 and May 28, 1999 as recorded by Albanian border guards. Magnitudes are on a log scale because of the large variation. The red "+" is the location of the Morina border crossing point with Albania.

Figure 2: Locations and Magnitudes of Albanian Civilians Killed in Kosovo



The locations and magnitudes of Albanians civilians killed in Kosovo between March 28, 1999 and May 28, 1999. Magnitudes are on a log scale because of the large variation. The red "+" is the location of the Morina border crossing point with Albania.

6 Methods

The theory of refugee cascades implies that refugees may positively affect the number civilians who flee from all other villages they encounter on their way to the border. However, identification of this global spillover effect is not possible. Instead, we can estimate a local spillover effect of fleeing only on neighboring villages. This local spillover effect must be smaller than the global spillover effect, so if a local spillover effect of fleeing exists, it provides evidence in favor of the theory.

Estimating the local causal effect of fleeing is challenging. OLS is biased and inconsistent in the presence of a spatial autoregressive lag (Anselin, 1988, pp. 2) and in many cases, identification of the parameters in spillover models is not possible (Manski, 1993, 2000). There are two identification problems cited by Manski (1993) that need to be overcome in order to estimate the spatial spillover effect of refugees fleeing. The first is disentangling the spillover effect of fleeing from spatially correlated omitted variables. Second, even in the absence of spatially correlated omitted variables, the simultaneous actions of neighboring villages affecting each other makes distinguishing the spillover effect of fleeing and the spillover effect of violence difficult.[11]

This suggests the linear model

$$y_{i,t} = \alpha + \beta \frac{\sum_{j \in N_i}}{n_i} y_{j,t} + \gamma x_i + \delta \frac{\sum_{j \in N_i}}{n_i} x_{j,t} + \epsilon_{i,t}, \tag{1}$$

where β represents the spillover effect of fleeing, γ is the local effect of violence and δ is the spillover effect of violence.

Using the spatial proximity of the villages to construct a matrix of spatial neighbors, W, equation 1 can be rewritten in matrix notation as

$$y_t = \alpha\iota + \beta W y_t + X_t\gamma + \delta W X_t + \epsilon_t \tag{2}$$

where ι is a vector of ones. This is similar to a spatial autoregressive model (SAR), $y = \beta W y + X\gamma + \epsilon$ (Cliff and Ord, 1981), with the addition of the spillover effect of violence. This model relaxes the assumption in a SAR that all spillover effects operate through the dependent variable. If civilians find out about violence in neighboring villages and this affects their probability of separately from the effect of refugees fleeing from the neighboring village a SAR model could find a spillover effect of fleeing when there was only a spillover effect of violence.

The first problem is that unobserved factors may affect villages that are near each other. Factors such as the course of the conflict, the economic situation are almost certainly spatially correlated. Residents in neighboring villages are likely to be more similar to each other than they are to residents of villages far away. Because of these spatially correlated omitted variables, if there is a correlation between fleeing from nearby villages it is difficult to determine whether this is truly due to spatial spillover effects or simply because nearby places are similar.

The second problem, which Manski (1993) refers to as the reflection problem, is due to fleeing from neighboring villages affecting each other. If refugees fleeing from one village may be caused by fleeing from a neighboring village, but those refugees' flight could have been caused by fleeing from the first village, we cannot know which village's fleeing is the cause and which is the effect.

The features of the spatial data in this case allow identification of the causal effects of spatial spillovers. In the model considered by Manski (1993), individuals are partitioned into groups,

[11] For clarity I refer to the three effects identified by Manski (1993, 2000) by names specific to this case. He refers to endogenous, contextual, and correlated effects. He considers a network of individuals, rather than villages, and defines endogenous effects as "the propensity of an agent to behave in some way varies with the behavior of the group," contextual effects as "the propensity of an agent to behave in some way varies with exogenous characteristics of the group members," and correlated effects as "agents in the same group tend to behave similarly because they have similar individual characteristics or face similar institutional environments" (Manski, 2000, pp. 127) I refer to these three effects as the spillover effect of fleeing, the spillover effect of violence, and spatially correlated omitted variables.

with everyone in a group affected by the average outcome of everyone else in that group, including themselves. Moffitt (2001) excludes each individual's outcome from the group average, but spillover effect still are not identified. Because the actions of everyone in a group simultaneously affect the actions of everyone else in the group and nothing outside the group affects their actions, the spillover effects cannot be identified. In spatial data this is not generally the case. Neighbors of your neighbors may not be your neighbor. Bramoullé, Djebbari and Fortin (2009) and De Giorgi, Pellizzari and Redaelli (2010) show that in such cases, where there is intransitivity in the neighbor connections, identification of the spillover effects is possible using characteristics of neighbors' neighbors and neighbors' neighbors' neighbors as instrumental variables.

However, the refugees fleeing themselves cause additional problems because they could potentially affect anyone on the way to the border, violating the exclusion restriction. The example illustrated in figure 3 demonstrates the identification problem and a solution using the unidirectional nature of refugees fleeing from Kosovo to Albania for a simple map of three villages that are increasingly close to the border. In this example neighboring villages within distance d can be defined using the symmetric spatial proximity matrix W_s.

$$W_s = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \tag{3}$$

Violence in village 3, x_3 can potentially cause fleeing from village 3, as well as spillover effects on its neighbor, village 2. Refugees generally move in the direction of the border, so there could be spillover effects of fleeing and violence on these villages closer to the border if people in these villages encounter refugees from village 3 on their way to the border. This makes identifying the spillover effect of fleeing on neighbors in the direction of the border difficult. However, because refugees will not travel far in the direction away from the border, there will not be direct spillover effects of fleeing or violence in the direction away from the border from non-neighboring villages. Therefore, violence in village 3, x_3, can serve as an instrumental variable for the effect of fleeing from neighboring villages on fleeing from their neighboring villages that are in the direction away from the border, y_2 on y_1.

These neighbors in the direction of the border can be defined using an asymmetric spatial proximity matrix, W_b. For the simple example shown in 3, the spatial proximity matrix of neighboring villages in the direction of the border is

$$W_b = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{4}$$

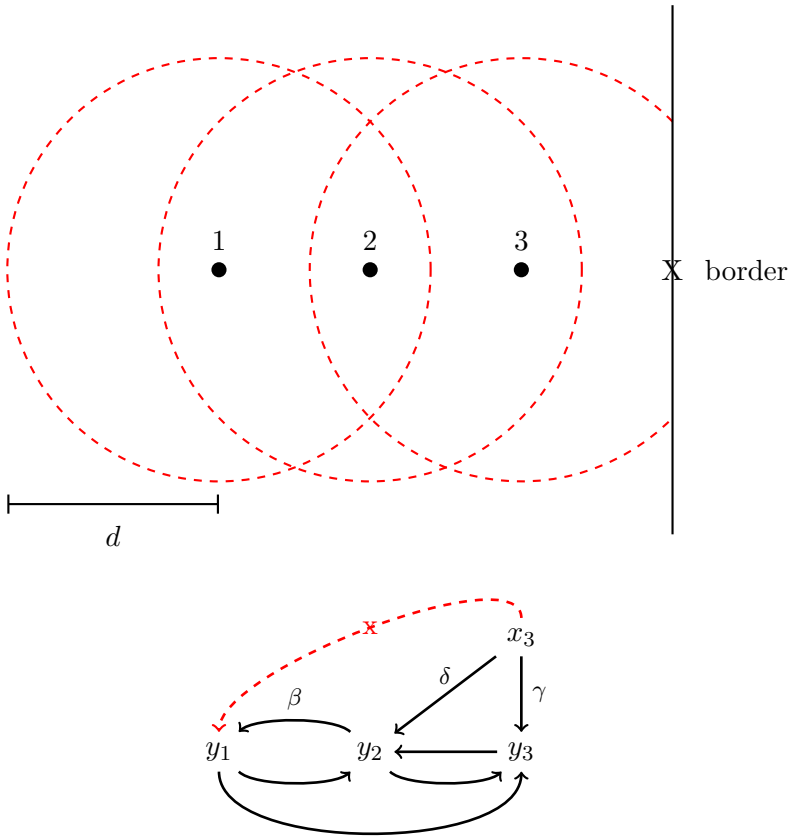
Then $W_b y$ is the number of refugees fleeing from neighboring villages that are closer to the border and $W_b^2 x$ is violence in the villages that neighbor neighboring villages in the direction of the border. $W_b^2 x$ can be used as an instrument for the spillover effect of fleeing, $W_b y$.

More generally, W^b is a spatial adjacency matrix that captures this relationship. In this matrix, village j is a neighbor of village i if they are near each other, within a threshold distance, d, and j does not pass i on the way to the borderis closer to the border than i.

$$W^b_{i,j} = \begin{cases} 1 & \text{if } \text{dist}(i,j) \leq d \text{ and } \text{dist}(i, \text{border}) > \text{dist}(j, \text{border}) \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

The spillover effect of fleeing only from neighbors in the direction towards the border can be identified by using $W_b^2 x$ as instrumental variables. Because the data on migration was collected from a single border crossing, we can use the spatial structure of the refugee migration to identify

Figure 3: Illustration of Identification Strategy



A simple example to illustrate the identification strategy. Above, a map of 3 villages, where village 1 is farthest from the border and village 3 is closest to the border. Dashed circles of radius d show which villages are neighboring other villages. The directed acyclic graph below shows the potential effects of violence in village 3, x_3 on fleeing from neighboring villages and from villages that neighbor neighboring villages. Because village 3 neighbors villages 2, there could be both spillover effects of fleeing and spillover effects of violence from village 3. Refugees flee in the direction of the border, so there could be spillover effects of fleeing and violence on these villages closer to the border if people in these villages encounter refugees from village 3 on their way to the border. This makes identifying the spillover effect of fleeing on neighbors in the direction of the border difficult. However, because refugees will not travel in the direction away from the border, there will not be direct spillover effects of fleeing or violence in the direction away from the border from non-neighboring villages. Therefore, violence in village 3, x_3, can serve as an instrumental variable for the effect of fleeing from neighboring villages on fleeing from their neighboring villages that are in the direction away from the border, y_2 on y_1.

the spillover effect of fleeing. Though no detailed information on the routes refugees took to get to the border is known, on average they must travel in the direction of the border.

Because of considerations of danger and terrain, the initial path a refugee takes out of his village could be in any direction. Ultimately, as refugees travel toward the border, they need to travel largely in the direction of the border, but the best path for a refugee fearing for his safety and with imperfect information about the larger environment could involve traveling in a direction for the first several kilometers that is not in the direction of the border. Additionally, because of long standing social ties caused by village proximity, civilians know others from neighboring villages, but, given the lack of widespread car ownership, the likelihood of having friends or acquaintances in another village falls off rapidly with distance. Before becoming refugees, there is no reason to believe there would be a bias in local social connections in the direction of the border—which was rarely crossed prior to the conflict. Fleeing refugees could deliberately warn friends in neighboring villages, which could be in any geographic direction, as they leave.

The identifying assumption is that while both refugees fleeing from neighboring villages and violence in neighboring villages can affect fleeing from your village directly, refugees fleeing from neighboring villages and violence in the neighboring villages of your neighboring villages which are not also neighboring your village can only affect fleeing from your village through their affect on your neighbors. In the particular case of Kosovo this assumption is valid in some geographical directions and not in others.

This identification strategy biases the results against finding a spillover effect of fleeing because refugees are probably more likely to initially flee in the direction of the border. So, even if the next village is only a few kilometers away, civilians living in villages in the direction of the border are more likely to learn from refugees nearby than civilians living in neighboring villages in the direction away from border.

There is still the potential for spatially correlated omitted variables, which could be time varying. For these to be problematic they would have to be regional, such that they affected a village, its neighbors and their neighbors. Since the dataset is a panel, I include municipality and day fixed effects.

There is still a need to define d, the threshold distance defining neighboring villages. It needs to be small enough that intra-village social networks plausibly exist or that refugees might travel near the neighboring village. The minimum distance such that all villages have at least one neighboring village in any direction is 6.3 km. I use five different values of d between 5 km and 25 km.

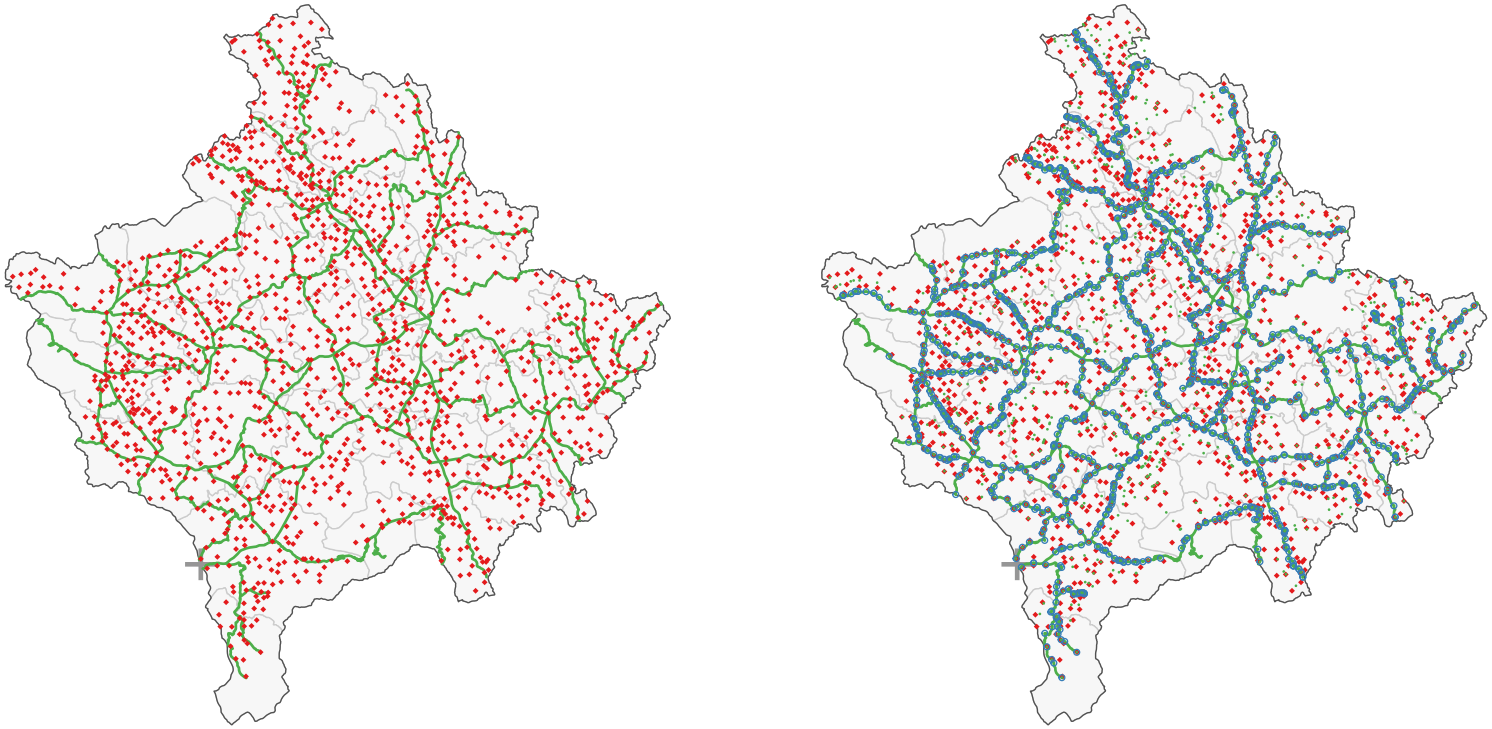
7 Fleeing Routes

The analysis in this paper hinges on estimating the routes refugees took from their home villages and the border. Fortunately the Albanian border crossing records record where refugees started their journeys and, because the data on refugee flight is all from the records of the Morina border post, I also know the end point of their journeys.

The vast majority of refugees traveled to the border on in vehicles on roads (). Using the major road network of Kosovo obtained from the Digital Chart of the World, which is derived from United States Defense Mapping Agency's (DMA) operational navigation chart (ONC) 1:1,000,000 scale paper map series produced by the US, Australia, Canada, and the United Kingdom. It is from 1992, prior to the war and there was not major road construction in Kosovo during the 1990s.

Using the road network, I estimate find the shortest route to the border on the road network for all villages. Figure 4 shows the highway network of Kosovo and all villages.

This also allows a different test of the theory: comparing refugee flows from nearby villages that are close to and farther from main roads. While individuals living in nearby villages experience similar risks, those that live near main roads are more likely to see and meet refugees fleeing, so spillover effects are expected to be larger than in villages farther from main roads.



(a) Settlements and the highway network

(b) Settlements, the highway network, and the nearest point on the highway network to each settlement

Figure 4: Proximity of settlements to highways

8 Results

Uncovering the causal spillover effect of refugees fleeing relies on the exclusion restriction that for village i on day t, refugees fleeing from the neighboring villages of neighboring villages in the direction of the border do not affect refugees fleeing from village i except through their effect on refugees fleeing from the village i's neighboring villages. This could be violated if an omitted variable causes both fleeing from village i and fleeing from the neighbors of neighbors of village i on the same day. Omitted variables that violate the exclusion restriction would have to be regional to affect both i and its neighbors. For example, if large troop movements in an area are observed by residents of village i, its neighboring villages, and the neighbors of those neighbors, the exclusion restriction could be violated.

I deal with this in several ways. First, I vary the distance between villages that are considered neighbors from 10 km to 30 km. If spatially correlated omitted variables are clustered at any particular distance, increasing the distance that encompasses neighbors beyond this will solve the problem. Increases in the threshold distance for neighboring villages requires ever more distant spatially correlated omitted variables to violate the exclusion restriction. In case there are local or larger regional, spatially correlated omitted variables, I include village and municipality fixed effects in different specifications. Municipalities are local regions encompassing multiple villages. If omitted variables, vary systematically at these regional levels, this will control for these regional effects. If omitted variables instead vary at the village level, village-level fixed effects are included. There is also the potential for time varying omitted variables. To account for this, I also include day fixed effects.

Table 3 presents regression results of models of the form in equation 2 to find the spillover effect of refugees fleeing. These control for local village measures of violence, civilians killed, KLA fighters killed, battles between the Yugoslav army and the KLA, Yugoslav army soldiers killed by the KLA, and NATO airstrikes, as well as the same measures of violence in neighboring villages and demographics in the local village an its neighbors. Village or municipality and time fixed effects are also included. The models are estimated using two stage least squares with the covariates of the neighbors of neighboring villages that are closer to the border, $W_b^2 x$ as instrumental variables.[12] Estimates for different threshold distances for neighboring villages are shown between 10km and 30km. The F-tests to reject weak instruments surpasses the conventional threshold of 10 for all models. First stage results are shown in the appendix.

There is a large, positive spillover effect of fleeing. The spatial weights matrix is row standardized, so the spillover effect can be interpreted as the number of civilians induced to flee and become refugees when an average of one refugee flees from the neighboring villages that are nearer to the border. This result is consistent across different distance thresholds for neighbors. A neighboring village may itself have multiple neighbors, affecting them all this way, on average. Thus, fleeing from neighboring areas has an independent effect on fleeing from local or neighboring violence. This result implies that civilians living in war zones learn from the behavior of other people, who may or may not have access to more information than them, but who are in similar situations. This learning could happen directly, because friends of acquaintances in neighboring villages tell them that they are leaving, or indirectly, when civilians observe that their neighbors have left. This is consistent with the theory of cascades of refugees—that civilians living in conflict zones without assess to good information on their local risk of victimization infer from the decisions of other civilians to flee that the risk must be higher, magnifying the effect of actual violence. It also implies that the location of violence may matter for the number of refugees that result from the conflict.

An additional testable implication of the theory of refugee cascades is that civilians living in villages closer to major roads will be more likely to flee because others have fled than those who live in villages farther from roads. This is for two reasons. First, access to roads increases the size of an individual's social network before the war, so during the war there are more people

[12] Kelejian, Prucha et al. (1997) and Kelejian and Prucha (1998) first suggest a spatial two-stage least squares estimator with the linearly independent set of instrumental variables $[W^2x, W^3x, W^4x \ldots]$ to instrument for the spillover effect of refugees fleeing from neighboring villages closer to the border, $W_b y$.

Table 3: Instrumental Variables Regression Results using Different Neighbor Distances

	Neighbors within					
	10 km		20 km		30 km	
Spillover effect of fleeing						
Refugees	1.155**	1.070***	0.976***	0.565**	1.438***	1.251***
(neighboring villages)	(0.513)	(0.391)	(0.316)	(0.286)	(0.360)	(0.356)
Local effect of violence						
Civilians killed	0.270	0.556***	0.299*	0.589***	0.306*	0.598***
	(0.186)	(0.187)	(0.180)	(0.182)	(0.180)	(0.183)
KLA killed	2.308**	2.662**	2.114*	2.475**	2.115*	2.445**
	(1.109)	(1.107)	(1.084)	(1.081)	(1.084)	(1.083)
Battles	−0.004	0.038	−0.007	−0.071	−0.043	−0.055
	(0.871)	(0.892)	(0.607)	(0.619)	(0.537)	(0.549)
Army killed	−0.478	−0.847	0.060	−0.019	0.394	0.212
	(0.929)	(0.939)	(0.629)	(0.640)	(0.553)	(0.565)
NATO airstrike	1.635	−0.273	0.400	−0.289	1.091	0.227
	(1.409)	(1.408)	(1.010)	(1.018)	(0.880)	(0.890)
Local controls						
Population		0.003***		0.004***		0.004***
		(0.001)		(0.001)		(0.001)
Serbs		−0.007***		−0.007***		−0.007***
		(0.001)		(0.001)		(0.001)
Albanians		−0.001		−0.001		−0.001
		(0.001)		(0.001)		(0.001)
Distance to border		−0.0001**		−0.0001**		−0.0001*
		(0.0001)		(0.00004)		(0.00004)
Spillover effect of violence						
Civilians killed	−0.712	−0.902	2.605	3.562*	1.129	1.717
(neighboring villages)	(1.421)	(1.286)	(2.027)	(2.024)	(3.069)	(3.119)
KLA killed	−7.091	−7.919	−9.323	−2.085	−22.981	−21.723
(neighboring villages)	(5.822)	(5.790)	(13.360)	(13.158)	(20.858)	(21.008)
Battles	0.369	0.192	−0.264	−0.636	0.110	−0.446
(neighboring villages)	(1.164)	(1.120)	(1.230)	(1.232)	(1.515)	(1.528)
Army killed	0.672	1.219	−0.130	0.306	0.039	0.412
(neighboring villages)	(1.248)	(1.238)	(1.174)	(1.187)	(1.396)	(1.425)
NATO airstrikes	−2.262	0.376	−0.210	0.817	−2.289	−0.014
(neighboring villages)	(1.806)	(1.762)	(1.728)	(1.733)	(2.033)	(2.044)
Neighbor controls						
Population		−0.004		0.008		−0.006
(neighboring villages)		(0.003)		(0.007)		(0.010)
Serbs		0.0003		−0.013		0.008
(neighboring villages)		(0.006)		(0.011)		(0.019)
Albanians		0.003		−0.006		0.005
(neighboring villages)		(0.003)		(0.008)		(0.011)
Distance to border		0.0001		0.0001		0.0001
(neighboring villages)		(0.0001)		(0.0001)		(0.0001)
Day fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Village fixed effects	Yes	No	Yes	No	Yes	No
Municipality fixed effects	No	Yes	No	Yes	No	Yes
Observations	89,404	89,404	89,404	89,404	89,404	89,404

Note: Standard errors clustered at the municipality level are shown, *p<0.1; **p<0.05; ***p<0.01

Table 4: Instrumental Variables Regression on Villages Near or Far from Highways

	Neighbors within					
	10 km		20 km		30 km	
	near	far	near	far	near	far
Spillover effect of fleeing						
Refugees (neighboring villages)	1.329*	−0.007	1.712***	−0.009	2.619***	0.014
	(0.756)	(0.015)	(0.547)	(0.009)	(0.664)	(0.009)
Local effect of violence						
Civilians killed	0.335	−0.002	0.413	−0.002	0.439	−0.002
	(0.300)	(0.006)	(0.288)	(0.006)	(0.288)	(0.006)
KLA killed	3.554**	−0.003	3.340*	0.001	3.296*	−0.003
	(1.774)	(0.033)	(1.752)	(0.033)	(1.756)	(0.033)
Battles	−0.231	0.005	−0.290	−0.002	−0.348	−0.008
	(1.830)	(0.020)	(1.271)	(0.014)	(1.163)	(0.012)
Army killed	−0.941	−0.098***	−0.038	−0.044**	0.563	−0.012
	(1.620)	(0.024)	(1.066)	(0.018)	(0.943)	(0.015)
NATO airstrike	3.005	−0.011	0.652	−0.001	1.847	−0.007
	(2.408)	(0.038)	(1.706)	(0.029)	(1.493)	(0.025)
Spillover effect of violence						
Civilians killed (neighboring villages)	0.966	−0.008	4.856	−0.040	0.941	−0.194**
	(2.033)	(0.046)	(3.433)	(0.057)	(5.199)	(0.087)
KLA killed (neighboring villages)	−9.540	0.058	−20.487	0.129	−52.981	−0.045
	(9.475)	(0.169)	(23.243)	(0.367)	(37.653)	(0.551)
Battles (neighboring villages)	−0.527	−0.021	−0.463	−0.012	0.530	0.059
	(2.716)	(0.025)	(2.312)	(0.031)	(2.642)	(0.041)
Army killed (neighboring villages)	1.665	0.165***	0.292	0.147***	0.212	0.082**
	(2.175)	(0.032)	(1.965)	(0.034)	(2.378)	(0.039)
NATO airstrikes (neighboring villages)	−3.887	−0.019	−0.859	−0.066	−4.635	−0.077
	(3.119)	(0.048)	(2.954)	(0.048)	(3.502)	(0.056)
Day fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Village fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Municipality fixed effects	No	No	No	No	No	No
Observations	52,638	36,766	52,638	36,766	52,638	36,766

Note: Standard errors clustered at the municipality level are shown, *p<0.1; **p<0.05; ***p<0.01

nearby who could potentially provide a warning. Second, proximity to roads provides more access to information by increasing the likelihood of seeing and meeting fleeing refugees. This leads to the hypothesis that the spillover effect will be stronger in villages closer to major roads. To test this I partitioned the data into two subsets: villages nearer and farther than 1,786 meters from the nearest highways—the average distance of a village from a highway. The results are shown in Table 4. The spillover effect of fleeing is large and statistically significant in villages near highways, but insignificant far from roads. The magnitudes are higher in the villages near highways than in the main results from Table 3.

There is a possible concern about the inclusion of variables measuring Yugoslav Army soldiers killed by the KLA, battles between the KLA and the Yugoslav Army, and NATO airstrikes, which are only measured at the municipality-day level, not the village-day level. Similar models excluding these municipality-level measures of violence yield substantively similar results and are shown in the appendix. All results include standard errors clustered at the municipality level because of the potential for locally correlated error terms.

9 Discussion and Conclusion

Based on a novel identification strategy using the neighbors of neighboring villages and the routes of refugees fleeing on roads in Kosovo, this paper finds that exposure to refugees from neighboring villages significantly increases the number of refugees that flee. This result is consistent with the proposed theoretical argument, that civilians living in conflict update their beliefs about the likelihood of victimization based on observing the decisions to flee of neighbors and that this can cause cascades of refugees streaming toward the border, in spite of little or no change in the probability of being killed.

By systematically identifying the spillover effect of fleeing, this study provides novel insight into how refugee flows can quickly grow and contributes to the growing body of research on the

causes of civilian displacement in civil war. In the information-poor environment of civil wars, the actions of others can have a substantial impact on an individual's own perception of the risk of violence and therefore on her decision to flee and become a refugee. If this civilian flees, she can further impact the information and decision making processes of other civilians, creating a refugee cascade caused by the information transmitted to other civilians by a stream of refugees fleeing. This can help explain why there is such variation in the numbers of refugees that result from wars with similar levels of violence.

Although the analysis in this paper is based on one conflict, the results can be applied more broadly. When refugees fleeing can affect other civilians' beliefs about risk and decision making, the spatial and temporal distribution of violence matters. The same amount of violence in different places will result in different numbers of refugees because cascades of refugees generated by information spillovers will differ.

A deeper understanding of why civilians flee and become refugees or stay in war zones is vital to preparing for and responding to current and future refugee crises. Expanding the reasons for refugee flows beyond local violence and economic factors and focusing on the role of information on the actions, not just of combatants, but civilians as well is a first step.

References

Adhikari, Prakash. 2012. "The Plight of the Forgotten Ones: Civil War and Forced Migration." *International Studies Quarterly* 56(3):590–606.

Adhikari, Prakash. 2013. "Conflict-Induced Displacement, Understanding the Causes of Flight." *American Journal of Political Science* 57(1):82–89.

Anselin, L. 1988. *Spatial Econometrics: Methods and Models*. Vol. 4 Springer Science & Business Media.

Apodaca, Clair. 1998. "Human Rights Abuses: Precursor to Refugee Flight?" *Journal of Refugee Studies* 11(1):80–93.

Arkin, William M. 2000. *Civilian Deaths in the NATO Air Campaign*. Vol. 12 Human Rights Watch.

Balcells, Laia and Abbey Steele. 2012. Warfare, Political Identities, and Displacement in Spain and Colombia. Technical report.

Ball, Patrick. 2000. *Policy or Panic?: The Flight of Ethnic Albanians from Kosovo, March-May 1999*. American Association for the Advancement of Science.

Ball, Patrick, Wendy Betts, Fritz Scheuren, Jana Dudukovich and Jana Asher. 2002. "Killings and Refugee flow in Kosovo March-June 1999." *Report to the International Criminal Tribunal for the former Yugoslavia* .

Banerjee, Abhijit V. 1992. "A Simple Model of Herd Behavior." *The Quarterly Journal of Economics* pp. 797–817.

Bieber, Florian. 2015. "The Construction of National Identity and its Challenges in Post-Yugoslav Censuses." *Social Science Quarterly* 96(3):873–903.

Bikhchandani, Sushil, David Hirshleifer and Ivo Welch. 1992. "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades." *Journal of Political Economy* pp. 992–1026.

Bikhchandani, Sushil, David Hirshleifer and Ivo Welch. 1998. "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades." *The Journal of Economic Perspectives* 12(3):151–170.

Bohra-Mishra, Pratikshya and Douglas S Massey. 2011. "Individual Decisions to Migrate during Civil Conflict." *Demography* 48(2):401–424.

Bramoullé, Yann, Habiba Djebbari and Bernard Fortin. 2009. "Identification of Peer Effects through Social Networks." *Journal of econometrics* 150(1):41–55.

Cliff, Andrew David and J Keith Ord. 1981. *Spatial Processes: Models & Applications*. Taylor & Francis.

Dancygier, Rafaela M and David D Laitin. 2014. "Immigration into Europe: Economic discrimination, violence, and public policy." *Annual Review of Political Science* 17:43–64.

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.

De Giorgi, Giacomo, Michele Pellizzari and Silvia Redaelli. 2010. "Identification of Social Interactions through Partially Overlapping Peer Groups." *American Economic Journal: Applied Economics* 2(2):241–275.

Edwards, Scott. 2009. *The Chaos of Forced Migration: A Means of Modeling Complexity for Humanitarian Ends*. VDM Publishing.

Engel, Stefanie and Ana María Ibáñez. 2007. "Displacement Due to Violence in Colombia: A Household-Level Analysis." *Economic Development and Cultural Change* 55(2):335–365.

Fischer, Peter A, Reiner Martin and Thomas Straubhaar. 1997. Should I Stay or Should I Go? In *International Migration, Immobility and Development*, ed. Kristof Tamas Tomas Hammar, Grete Brochmann and Thomas Faist. Berg: Oxford and New York pp. 49–90.

Gibney, Mark, Clair Apodaca and James McCann. 1996. Refugee Flows, the Internally Displaced, and Political Violence (1980-1993): An Exploratory Analysis. In *Whither Refugee? The Refugee Crisis: Problems and Solutions*, ed. Alex Schmid. Ploom pp. 45–66.

Hirschmann, Albert O. 1970. "Exit, Voice and Loyalty." *Responses to Decline in Firms, Organizations, and States. Cambridge* .

Iacopino, Vincent, Martina W Frank, Heidi M Bauer, Allen S Keller, Sheri L Fink, Doug Ford, Daniel J Pallin and Ronald Waldman. 2001. "A Population-based Assessment of Human Rights Abuses Committed Against Ethnic Albanian Refugees from Kosovo." *American Journal of Public Health* 91(12):2013–2018.

Ibáñez, Ana María and Carlos Eduardo Vélez. 2008. "Civil Conflict and Forced Migration: The micro determinants and welfare losses of displacement in Colombia." *World Development* 36(4):659–676.

Jonassohn, Kurt. 1993. "Famine, genocide, and refugees." *Society* 30(6):72–76.

Kalyvas, Stathis. 2006. *The Logic of Violence in Civil War*. Cambridge University Press.

Kalyvas, Stathis N. 1999. "Wanton and senseless? The logic of massacres in Algeria." *Rationality and Society* 11(3):243–285.

Kelejian, Harry H and Ingmar R Prucha. 1998. "A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances." *The Journal of Real Estate Finance and Economics* 17(1):99–121.

Kelejian, Harry H, Ingmar R Prucha et al. 1997. "Estimation of Spatial Regression Models with Autoregressive Errors by Two Stage Least Squares Procedures: A serious problem." *International Regional Science Review* 20(1):103–112.

King, Gary, Robert O Keohane and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton University Press.

Klingner, Jeff and Romesh Silva. 2013. Combining Found Data and Surveys to Measure Conflict Mortality. In *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict*, ed. Taylor B Seybolt, Jay D Aronson and Baruch Fischhoff. Oxford University Press pp. 147–164.

Lozano-Gracia, Nancy, Gianfranco Piras, Ana Maria Ibáñez and Geoffrey JD Hewings. 2010. "The journey to safety: conflict-driven migration flows in Colombia." *International Regional Science Review* 33(2):157–180.

Manski, Charles F. 1993. "Identification of endogenous social effects: The reflection problem." *The review of economic studies* 60(3):531–542.

Manski, Charles F. 2000. "Economic Analysis of Social Interactions." *The Journal of Economic Perspectives* 14(3):115–136.

Melander, Erik and Magnus Öberg. 2006. "Time to go? Duration dependence in forced migration." *International Interactions* 32(2):129–152.

Melander, Erik and Magnus Öberg. 2007. "The Threat of Violence and Forced Migration: Geographical Scope Trumps Intensity of Fighting." *Civil Wars* 9(2):156–173.

Moffitt, Robert A. 2001. Policy Interventions, Low-level Equilibria, and Social Interactions. In *Social Dynamics*, ed. Steven Durlauf and Peyton Young. Vol. 4 MIT Press pp. 6–17.

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. 2006. "Refugee or internally displaced person? To where should one flee?" *Comparative Political Studies* 39(5):599–622.

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.

Nordstrom, Carolyn. 1997. *A Different Kind of war Story*. University of Pennsylvania Press.

Rummel, Rudolph J. 1994. "Power, Genocide and Mass murder." *Journal of Peace Research* 31(1):1–10.

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.

Schmeidl, Susanne. 1997. "Exploring the causes of forced migration: A pooled time-series analysis, 1971-1990." *Social Science Quarterly* pp. 284–308.

Snyder, Richard. 2001. "Scaling down: The subnational comparative method." *Studies in comparative international development* 36(1):93–110.

Spitzer, Yannay. 2015. "Pogroms, Networks, and Migration: The Jewish Migration from the Russian Empire to the United States, 1881–1914." *Working paper* .

Steele, Abbey. 2011. "Electing Displacement: Political Cleansing in Apartadó, Colombia." *Journal of Conflict Resolution* 55(3):423–445.

Walter, Barbara F. 2009. "Bargaining failures and civil war." *Annual Review of Political Science* 12:243–261.

Weiner, Myron. 1998. "The Clash of Norms: Dilemmas in refugee policies." *Journal of Refugee Studies* 11(4):433–453.

Wiesner, Louis A. 1988. *Victims and Survivors: Displaced Persons and Other War Victims in Viet-Nam, 1954-1975*. Greenwood Pub Group.

A Appendix

Table A.1: First stages of models in table 3

	Neighbors within					
	10 km		20 km		30 km	
Instruments						
Civilians killed (neighbors of neighboring villages)	0.542 (0.383)	0.419 (0.384)	1.215** (0.492)	1.274*** (0.489)	8.353*** (0.488)	8.120*** (0.484)
KLA killed (neighbors of neighboring villages)	21.194*** (2.533)	19.728*** (2.528)	79.314*** (3.068)	78.819*** (3.040)	77.033*** (3.262)	75.999*** (3.220)
Battles (neighbors of neighboring villages)	−0.256 (0.440)	−0.770* (0.443)	−2.437*** (0.327)	−2.876*** (0.327)	−1.633*** (0.277)	−1.963*** (0.275)
Army killed (neighbors of neighboring villages)	−0.418 (0.565)	0.157 (0.564)	−1.380*** (0.332)	−1.065*** (0.330)	−2.496*** (0.304)	−2.454*** (0.301)
NATO airstrikes (neighbors of neighboring villages)	−0.295 (0.808)	2.621*** (0.763)	1.749*** (0.490)	3.337*** (0.476)	4.047*** (0.453)	5.387*** (0.439)
Population (neighbors of neighboring villages)		−0.006*** (0.001)		0.015*** (0.001)		0.011*** (0.003)
Serbs (neighbors of neighboring villages)		0.006*** (0.002)		−0.014*** (0.003)		−0.024*** (0.004)
Albanians (neighbors of neighboring villages)		0.006*** (0.001)		−0.018*** (0.002)		−0.012*** (0.003)
Distance to border (neighbors of neighboring villages)		−0.0002*** (0.00004)		−0.0001*** (0.00003)		0.00004** (0.00002)
Controls						
Civilians killed	0.029 (0.038)	0.029 (0.038)	0.006 (0.019)	0.001 (0.019)	−0.018 (0.013)	−0.024* (0.013)
KLA killed	−0.458** (0.227)	−0.479** (0.224)	−0.331*** (0.116)	−0.304*** (0.113)	−0.097 (0.078)	−0.101 (0.076)
Battles	−0.090 (0.179)	−0.367** (0.180)	−0.011 (0.064)	−0.038 (0.064)	−0.065 (0.040)	−0.058 (0.040)
Army killed	0.135 (0.186)	0.150 (0.186)	0.191*** (0.067)	0.186*** (0.067)	0.077* (0.041)	0.087** (0.041)
NATO airstrikes	−0.087 (0.281)	0.548** (0.274)	0.036 (0.107)	0.129 (0.106)	−0.346*** (0.064)	−0.307*** (0.063)
Population		0.0002* (0.0001)		0.00001 (0.0001)		−0.0001** (0.00005)
Serbs		−0.0001 (0.0002)		−0.0001 (0.0001)		0.0001 (0.0001)
Albanians		−0.0003** (0.0002)		−0.00001 (0.0001)		0.0001*** (0.0001)
Distance to border		0.00001 (0.00001)		0.00001** (0.00000)		0.00000 (0.00000)
Civilians killed (neighboring villages)	1.722*** (0.268)	1.993*** (0.268)	1.410*** (0.325)	1.427*** (0.322)	−0.873*** (0.308)	−0.666** (0.305)
KLA killed (neighboring villages)	−6.134*** (1.578)	−5.372*** (1.568)	−14.252*** (1.821)	−13.968*** (1.798)	−3.596* (1.905)	−3.208* (1.871)
Battles (neighboring villages)	−0.560 (0.454)	0.263 (0.456)	0.196 (0.232)	0.632*** (0.231)	−0.599*** (0.179)	−0.310* (0.178)
Army killed (neighboring villages)	0.839 (0.556)	0.280 (0.554)	1.590*** (0.259)	1.343*** (0.257)	1.544*** (0.205)	1.513*** (0.203)
NATO airstrikes (neighboring villages)	0.737 (0.800)	−2.907*** (0.742)	−1.190*** (0.406)	−2.730*** (0.390)	−1.719*** (0.324)	−2.822*** (0.311)
Population (neighboring villages)		0.006*** (0.001)		0.001 (0.001)		0.005*** (0.001)
Serbs (neighboring villages)		−0.008*** (0.001)		−0.007*** (0.001)		−0.007*** (0.002)
Albanians (neighboring villages)		−0.004*** (0.001)		0.002* (0.001)		−0.004*** (0.001)
Distance to border (neighboring villages)		0.0002*** (0.00004)		0.00005** (0.00002)		−0.00004** (0.00001)
Day fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Village fixed effects	Yes	No	Yes	No	Yes	No
Municipality fixed effects	No	Yes	No	Yes	No	Yes
Observations	89,404	89,404	89,404	89,404	89,404	89,404
R^2	0.005	0.088	0.021	0.196	0.033	0.329
Adjusted R^2	−0.013	0.087	0.004	0.195	0.017	0.328
F Statistic	26.680***	73.954***	123.111***	186.545***	201.170***	374.086***

Note: Standard errors clustered at the municipality level are shown, *p<0.1; **p<0.05; ***p<0.01

Table A.2: Instrumental Variables Regression Results Excluding Municipality-level Covariates

	Neighbors within					
	10 km		20 km		30 km	
Spillover effect of fleeing						
Refugees (neighboring villages)	1.190**	0.927**	1.010***	0.607**	1.488***	1.453***
	(0.508)	(0.402)	(0.332)	(0.303)	(0.391)	(0.396)
Local effect of violence						
Civilians killed	0.267	0.560***	0.299*	0.589***	0.308*	0.598***
	(0.186)	(0.186)	(0.180)	(0.182)	(0.181)	(0.183)
KLA killed	2.322**	2.669**	2.106*	2.464**	2.127**	2.429**
	(1.111)	(1.100)	(1.084)	(1.081)	(1.084)	(1.085)
Local controls						
Population		0.003***		0.004***		0.004***
		(0.001)		(0.001)		(0.001)
Serbs		−0.007***		−0.007***		−0.007***
		(0.001)		(0.001)		(0.001)
Albanians		−0.001		−0.001		−0.001
		(0.001)		(0.001)		(0.001)
Distance to border		−0.0001**		−0.0001**		−0.0001*
		(0.0001)		(0.00004)		(0.00004)
Spillover effect of violence						
Civilians killed (neighboring villages)	−0.821	−0.565	2.561	3.506*	0.933	1.009
	(1.422)	(1.297)	(2.045)	(2.038)	(3.129)	(3.195)
KLA killed (neighboring villages)	−7.118	−7.220	−10.308	−3.188	−23.186	−26.780
	(5.856)	(5.779)	(13.525)	(13.298)	(20.849)	(21.180)
Neighbor controls						
Population (neighboring villages)		−0.003		0.007		−0.008
		(0.003)		(0.007)		(0.010)
Serbs (neighboring villages)		−0.001		−0.013		0.011
		(0.006)		(0.012)		(0.019)
Albanians (neighboring villages)		0.003		−0.006		0.007
		(0.003)		(0.008)		(0.011)
Distance to border (neighboring villages)		0.0001		0.0001		0.0001
		(0.0001)		(0.0001)		(0.0001)
Day fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Village fixed effects	Yes	No	Yes	No	Yes	No
Municipality fixed effects	No	Yes	No	Yes	No	Yes
Observations	89,404	89,404	89,404	89,404	89,404	89,404

Note: Standard errors clustered at the municipality level are shown, *p<0.1; **p<0.05; ***p<0.01

Table A.3: OLS Regression Results using different Neighbor Distances

	Neighbors within					
	10 km		20 km		30 km	
Spillover effect of fleeing						
Refugees (neighboring villages)	−0.048***	−0.086***	0.004	−0.039	−0.070	−0.132***
	(0.016)	(0.016)	(0.032)	(0.032)	(0.046)	(0.047)
Local effect of violence						
Civilians killed	0.322*	0.601***	0.319*	0.599***	0.326*	0.607***
	(0.179)	(0.182)	(0.179)	(0.182)	(0.179)	(0.182)
KLA killed	2.270**	2.585**	2.265**	2.576**	2.317**	2.614**
	(1.077)	(1.077)	(1.077)	(1.077)	(1.077)	(1.077)
Battles	−0.061	−0.301	−0.090	−0.146	−0.256	−0.267
	(0.845)	(0.861)	(0.603)	(0.616)	(0.531)	(0.543)
Army killed	−0.310	−0.650	0.171	0.036	0.434	0.260
	(0.899)	(0.912)	(0.625)	(0.638)	(0.550)	(0.562)
NATO airstrike	1.518	0.577	0.501	−0.190	0.719	0.009
	(1.367)	(1.342)	(1.004)	(1.015)	(0.870)	(0.884)
Local controls						
Population		0.003***		0.004***		0.004***
		(0.001)		(0.001)		(0.001)
Serbs		−0.007***		−0.008***		−0.007***
		(0.001)		(0.001)		(0.001)
Albanians		−0.001		−0.001		−0.001
		(0.001)		(0.001)		(0.001)
Distance to border		−0.0001**		−0.0001**		−0.0001*
		(0.00005)		(0.00004)		(0.00004)
Spillover effect of violence						
Civilians killed (neighboring villages)	1.875**	1.790**	5.122***	5.127***	6.385**	6.635**
	(0.870)	(0.884)	(1.847)	(1.881)	(2.789)	(2.843)
KLA killed (neighboring villages)	−2.960	−3.622	9.067	8.830	20.086	17.584
	(5.388)	(5.457)	(11.902)	(12.092)	(18.094)	(18.377)
Battles (neighboring villages)	−0.554	−0.231	−1.184	−1.086	−1.727	−1.926
	(1.063)	(1.082)	(1.187)	(1.211)	(1.442)	(1.474)
Army killed (neighboring villages)	1.355	1.736	0.820	0.923	0.334	0.712
	(1.178)	(1.193)	(1.127)	(1.149)	(1.386)	(1.416)
NATO airstrike (neighboring villages)	−1.633	−0.716	−0.271	0.602	−1.224	0.561
	(1.734)	(1.678)	(1.718)	(1.726)	(2.005)	(2.029)
Neighbor controls						
Population (neighboring villages)		0.001		0.012**		0.007
		(0.003)		(0.006)		(0.009)
Serbs (neighboring villages)		−0.007		−0.022**		−0.012
		(0.005)		(0.011)		(0.018)
Albanians (neighboring villages)		0.001		−0.011		−0.008
		(0.003)		(0.007)		(0.010)
Distance to border (neighboring villages)		0.0001*		0.0001		0.00004
		(0.0001)		(0.0001)		(0.0001)
Day fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Village fixed effects	Yes	No	Yes	No	Yes	No
Municipality fixed effects	No	Yes	No	Yes	No	Yes
Observations	89,404	89,404	89,404	89,404	89,404	89,404

Note: Standard errors clustered at the municipality level are shown, *p<0.1; **p<0.05; ***p<0.01