A spatial model of internal displacement and forced migration

Jon Echevarria and Javier Gardeazabal*
University of the Basque Country UPV/EHU†

December 6, 2019

Abstract

This article develops a spatial model of internal and external forced migration. We propose a model reminiscent of Hoteling's spatial model in economics and Schelling's model of segregation. Conflict is modeled as a shock that hits a country at certain location and generates displacement of people located near the shock's location. Some displaced people cross a border, thus becoming refugees, while others remain as Internally Displaced Persons (IDPs). The model delivers predictions about how the fractions of a country's population that become refugees and IDPs ought to be related with the intensity of the shock, country size, terrain ruggedness and the degree of geographical proximity of the country with respect to the rest of the world. The predictions of the model are then tested against real data using a panel of 200 countries covering the period 1960-2016. The empirical evidence is broadly in line with the predictions of the model.

KEYWORDS: Internal displacement; forced migration; spatial model; conflict

WORD COUNT: 9,983

*Thanks to Daniel Arce, María Paz Espinosa, Irene Olloqui, Clint Peinhardt, James Piazza, Todd Sandler, Fernando Sanz, William Shughart, Thomas Zeitzoff, and participants at the 2018 Political Violence and Policy Conference at the University of Texas at Dallas, the 2018 Australasian Meeting of the Econometric Society at Auckland University of Technology and the 2018 Meeting of the Spanish Economic Association (SAEe 2018) held at Getafe Campus at the Universidad Carlos III de Madrid. Jon Echevarria acknowledges financial support from the Basque Government (PRE-2015-1-0073). Javier Gardeazabal acknowledges financial support from the Basque Government (IT783-13) and the Spanish Ministry of Economics and Competitiveness (Mineco-FEDER ECO2015-64467-R).

†Mailing address: Departamento de Fundamentos del Análisis Económico II, Universidad del País Vasco / Euskal Herriko Unibertsitatea, Lehendakari Aguirre 83, 48015 Bilbao, Spain. E-mail: `jon.echevarria@ehu.eus` and `javier.gardeazabal@ehu.eus`

Forced displacement occurs when a group of people are obliged to leave their home location unwillingly. Instead, economic migrants choose to leave their home in search of economic opportunities. Forcibly displaced persons can be classified into several groups depending on the root cause of displacement and whether they cross an international border or not. This paper focuses on two groups of forcibly displaced persons: refugees and Internally Displaced Persons (IDPs).

According to the 1951 United Nations High Commissioner for Refugees (UNHCR) Refugee Convention and the 1967 Protocol (UNHCR 2011), a refugee is someone who "owing to a well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group or political opinion, is outside the country of his nationality, and is unable to, or owing to such fear, is unwilling to avail himself of the protection of that country." The 1969 Organization of African Unity Convention (UNHCR 2006) extended the definition to include those who escape from origin countries because of "acts of external aggression, occupation, domination by foreign powers or serious disturbances of public order."

The UNHCR defines IDPs as those "...who have been forced or obliged to flee or to leave their homes or places of habitual residence, in particular as a result of or in order to avoid the effects of armed conflict, situations of generalized violence, violations of human rights or natural or human-made disasters, and who have not crossed an internationally recognized State border" according to the original definition agreed upon at the United Nations in 1998, e.g. Kälin (2008).

According to these definitions, forced displacement can be caused by armed conflict, lack of civil liberties or political rights and natural disasters. However, internal and external displacement share a common cause: armed conflict. When we focus on displaced people who escape from armed conflict, the distinguishing feature between refugees and IDPs is whether they cross an international border. As refugees and IDPs have a common cause it seems logical to study both groups of people simultaneously. However, the literature on forced displacement is mainly divided into refugee studies on the one hand and internal displacement studies on the other.

In this paper, we develop a theoretical model where armed conflict is the common cause of internal and external forced migration. We propose a spatial model of internal displacement and forced migration which has no predecessor in the literature of forced migration. Our model is reminiscent of Hoteling's spatial market model (Hotelling, 1929) and Schelling's model of segregation (Schelling, 1971). As in Hotelling's model, population is uniformly distributed in a segment of the real line. As in Schelling's model, people change their location according to the dynamics of the model. The spatial model we develop represents armed conflict as a shock that takes place at a particular location and generates a migration flow. Some displaced people cross a border, thus becoming refugees, while others remain as IDPs. We abstract from other known causes of internal and external forced migration, such as lack of civil liberties or political rights, economic development and natural disasters.[1] The model predicts how the number of refugees and IDPs, as a fraction of a country's population, varies with the intensity of armed conflict and geographical covariates such as country size, orography and distance to other countries. In our model, armed conflict has a causal effect on displacement, and the geographical variables exert an effect modification on displacement.

One of the effect modifiers we study is physical country size, i.e. area. Basic intuition indicates that the area of a country must necessarily be a determinant in the fraction of the population affected by a conflict of a particular magnitude. For instance, in a very large country, a conflict might generate a certain amount of internal displacement of affected people to other parts of the country and no refugees. However, a conflict of similar magnitude in a sufficiently smaller country would generate an outflow or refugees to neighboring countries. A second geographical covariate of interest is the orography. Intuition suggests that displacement should be easier the flatter and obstacle-free a country is, while displacement ought to be much more difficult in a highly mountainous country. Thus, as a result of a conflict of a particular magnitude, the rougher the orography, the less the displaced people ought to move, hence potentially generating a lower number of refugees. A third geographical variable analyzed is distance to neighboring countries. Intuition suggests that the more distant neighboring countries are, the lower the number of refugees while total displacement should be independent of the distance. Summarizing, in the model, armed conflict generates displacement while country size, orography and distance to other countries determine the division of displacement into refugees and IDPs. We develop a barebones model intentionally, with the aim of keeping the mathematical analysis as simple as possible. Due to its simplicity, the model allows for a number of extensions at the cost of more complicated analysis.

In addition, this paper contributes empirical evidence by testing the predictions of the model against real data. The empirics use a panel data set of 200 countries covering the period 1960-2016. This data set includes country of origin refugee and IDP counts, conflict and geographical data. We exploit the cross country variation in the geographical variables to analyze their effect modification. In addition, we also account for non-spatial variables which our model ignores, but the literature has found relevant, such as the lack of civil liberties or political rights and Gross Domestic Product (GDP) per capita. The evidence obtained is in line with the model's predictions.

The next section reviews the literature and indicates the contribution of this paper to it. Section 2 develops a spatial model of internal and external displacement. Section 3 describes the empirical methods, the data used and reports the empirical evidence. Section 4 sums up the main conclusions.

1 Literature review and our contribution

The literature on forced displacement is divided into two areas: refugee and IDP studies. A fundamental reason for this divide in the literature is data availability. Refugee destination countries keep track of asylum applications which generate data which can be used to analyze the choice of destination, e.g. Havinga and Böcker (1999), Neumayer (2005a) and Neumayer (2005b). Other studies analyze dyadic forced displacement flows between countries, e.g. Iqbal (2007), Moore and Shellman (2007), Barthel and Neumayer (2015) and Echevarria and Gardeazabal (2016).

A stream of the literature builds on the consequences of refugee inflows in the host country. Some explore the role of refugee flows in spreading conflict in the host country e.g. Salehyan and Gleditsch (2006), Salehyan (2008), Milton et al. (2013), Rüegger (2019), Fisk (2019) and Böhmelt et al. (2019). Another arm focuses on the effect of refugee inflows on the host country's contribution to Official Development Assistance (ODA) (e.g. Czaika, 2009), the UNHCR (e.g. Roper and Barria, 2010) or UN peacekeeping missions (e.g. Uzonyi, 2015).

All previous references use cross-country data (either longitudinal or dyadic) to analyze external displacement. However, a large number of articles are country case studies. Following Card's (1990) seminal article on the labor market consequences of a large immigration shock, a literature has emerged, e.g. Borjas (2017), Peri and Yasenov (2019, forthcoming). Similarly, large refugee inflows have prompted research analyzing their effect on the host country's food prices, wealth, labor market, health, education and other outcomes, e.g. Alix-Garcia and Saah (2009), Baez (2011), Tumen (2016), Ceritoglu et al. (2017), Esen and Oğuş Binatlı (2017), Akgündüz and Torun (2018) and Verme and Schuettler (2019).

Relatively less work has been done on the analysis of refugee outflows from origin countries. Davenport et al. (2003) analyzes the determinants of aggregate stocks of migrants, whether internal or external, from origin countries. More recently, Dreher et al. (2019) analyze the effect of aid receipts on total refugee outflows and also on flows to donor countries. Our paper analyzes refugee outflows, thus contributing to this less prolific arm of the literature.

When we turn to the literature on internal displacement, we find country case studies. As in refugee studies, there are quantitative assessments of the effect of displacement on host communities. The massive conflict-induced internal displacement in Colombia has been the focus of several articles. Engel and Ibáñez (2007) and Ibáñez and Vélez (2008) study the determinants of displacement. Morales (2018) studies the impact of displacement on wages in the host communities. Depetris-Chauvin and Santos (2018) analyze the effect of internal displacement on rental prices in host cities. McEniry et al. (2019) investigate the effect of exposure to displacement on older adults health. Displacement determinants in Indonesia are analyzed in Czaika and Kis-Katos (2009). Kondylis (2008) conducts a resettlement policy evaluation in Rwanda. Kondylis (2010) assesses the labor market effect of displacement in Bosnia and Herzegovina. Alix-Garcia et al. (2012) study the price responses to internal displacement and aid in Sudan. Alix-Garcia et al. (2013) estimate the effect of conflict-induced internal displacement on spatial changes in land use. The spread of conflict has also been studied in relation with internal displacement, e.g. Bohnet et al. (2018).

Other country case studies analyze not only conflict-originated displacement but also internal displacement caused by natural disasters and economic development e.g. Lanjouw et al. (2000), Muggah (2003) and Fernandes (2017). There are also a few policy analyses Goswami (2007), Lischer (2008), Crisp (2010) and Munive (2019), and also applications of the interview-research methodology, e.g. Ayata and Yukseker (2005).

However, to the best of our knowledge, there is not a single cross-country study on internal displacement. We carry out a cross country empirical analysis of forced internal displacement, thus contributing evidence on this previously unexplored angle.

Theoretical studies of displacement are scant in the literature. A notable exception is Czaika (2009) who models refugee migration decisions and studies the distribution of burdens from forced migration across countries. Therefore, our model is a contribution in a field with scarce production of theoretical models.

Furthermore, the model designed below builds a bridge between internal and external forced migration. As far as we know, simultaneous modeling of internal and external forced migration has not been attempted before.

Indeed, armed conflict generates displacement. Ball et al. (2002) is perhaps the single most clear example showing the relation between conflict and displacement. Their study of the 1998-1999 conflict in Kosovo shows how displacement counts mirror conflict intensity measured as the number of people killed. Our model takes the relation between conflict and displacement as granted and analyzes how geographical variables affect displacement, both internally and externally. The literature includes previous studies where geographical factors play a role in determining displacement flows. In particular, the influence of distance between countries has been incorporated into empirical models of dyadic refugee counts, either through spatial dependence as in Barthel and Neumayer (2015) or directly as in the gravity models of Iqbal (2007) or Echevarria and Gardeazabal (2016). However, the roles of country size and ruggedness of terrain have not been explored before.

2 A spatial model of forced migration

For expositional purposes, first we analyze a baseline version of the model with no spatial variables involved in subsection 2. This version of the model is the easiest possible and allows us to understand its mechanics. Then, subsection 2.1 develops a more sophisticated version of the model where we include the spatial variables and see how they interact with conflict in the determination of displacement.

The baseline model

This section develops a spatial model to explain conflict induced displacement. Three countries, A, B and C, align in the real line. Country A is located to the left of the the origin, country B is located in the interval $[0,1]$ and country C to the right of unity. Points 0 and 1 are the borders. We will focus on country B, whose population is uniformly distributed on the unit interval. A shock of size s hits country B at location l_s, a point in the unit interval. Figure 1 shows the assumptions made so far.

The shock affects citizens of country B forcing them to leave their home location if the benefits from staying are lower or equal than the costs. Remaining at their home location yields a benefit b for every citizen, while the cost of staying is

$$c(s, l_i, l_s) = s - |l_i - l_s|, \tag{1}$$

where s is the size of the shock, l_i, the location of individual i, and l_s is the location of the shock. In words, the cost of staying equals the size of the shock minus the distance from the location of the individual to the location of the shock. For simplicity, we normalize the benefit of staying to zero. Under these assumptions, individuals whose distance to the shock location is less than the size of the shock are displaced.

We assume the intensity of the shock, s, ranges from 0 to 1. Notice that a shock whose intensity is equal to 1 would affect the entire population. Therefore, it seems unnecessary to analyze larger shocks. We also assume that there is no international spillover of armed conflict and therefore the shock only affects citizens of country B. Hence, in those cases where $l_s - s$ is negative or $l_s + s$ is greater than unity, those who decide to move are in the intersection of $[l_s - s, l_s + s]$ and the unit interval. Depending on whether $l_s - s$ is positive or negative and $l_s + s$ is smaller or greater than unity, the proportion of displaced people varies. The four feasible cases are represented by the four regions depicted in Figure 2.

Region 1 corresponds to the case where $l_s - s \geq 0$ and $l_s + s \leq 1$ and represented in Figure 3. Under these conditions, the proportion of people displaced equals $2s$, the length of the segment $[l_s - s, l_s + s]$. Within region 2, and as shown in Figure 4, $l_s + s \leq 1$ and $l_s - s < 0$, so the shock affects everyone located to the left of l_s and displaced people are a fraction $l_s + s$ of the population. In Region 3, $l_s - s \geq 0$ and $l_s + s > 1$, so displaced people are a fraction $1 - (l_s - s)$ of the population, as captured by Figure 5. Finally, in Region 4, $l_s - s < 0$ and $l_s + s >$, so every person is affected and the proportion of displaced people equals 1. Therefore, total displacement, D, is defined over the four regions as follows

$$D = \begin{cases} 2s & if \quad l_s \geq s \quad \& \quad l_s \leq 1-s \\ l_s+s & if \quad l_s < s \quad \& \quad l_s \leq 1-s \\ 1-(l_s-s) & if \quad l_s \geq s \quad \& \quad l_s > 1-s \\ 1 & if \quad l_s < s \quad \& \quad l_s > 1-s \end{cases} \tag{2}$$

Notice that total displacement is a continuous function of the shock's intensity and location, and non-differentiable at the boundaries between regions.

Under the assumptions laid out so far, the model predicts the proportion of the population displaced as a function of shock intensity and location. With no further assumptions, the model gives no indication as to how far displaced persons move. For simplicity, we assume that displaced persons move a distance equal to the size of the shock, either leftward or rightward depending on which side of the shock location they are. This form of displacement is consistent with a story where at each point in the line to the left of $l_s - s$ or the right of $l_s + s$, a person not affected by the shock hosts one, and only one, displaced person. As a consequence of people's displacement, some might cross a border becoming refugees while others remain in the home country as IDPs. More precisely, those people whose distance to a border is less than the size of the shock become refugees. For the time being, we will assume that crossing a border is costless and relax this assumption in Section 2.1.

Figure 6 sorts (s, l_s)-pairs into nine regions depending of their position with respect to four lines. Region 1 is further described in Figure 7, where the shock affects everyone in interval $[l_s - s, l_s + s]$. Those in the interval $[l_s - s, l_s)$ move to $[l_s - 2s, l_s - s)$ and those initially in the interval $[l_s, l_s + s]$ move to $(l_s + s, l_s + 2s]$. Because $l_s > 2s$, no one leaves the country through the border with country A, and as $l_s < 1 - 2s$, no one crosses the border with country C. As a consequence, everyone affected by the shock remains internally displaced. Figure 8 represents Region 2 where, as in the previous case, no one reaches the border to Country A, while a fraction $l_s + 2s - 1$ of the population reaches Country C and become refugees. Figure 9 represents Region 3 where all people affected to the right side of l_s, the fraction $1 - l_s$ of the population, become refugees while those affected to the left of l_s remain as IDPs. The reader is invited to carry out the same sort of reasoning for the remaining regions to verify that those who become refugees represent a fraction R of the population given by the following expression,

$$
R = \begin{cases} 0 & if \quad 0 < l_s - 2s < l_s - s \;\; \& \;\; l_s + s < l_s + 2s < 1 \\ l_s + 2s - 1 & if \quad 0 < l_s - 2s < l_s - s \;\; \& \;\; l_s + s < 1 \leq l_s + 2s \\ 1 - l_s & if \quad 0 < l_s - 2s < l_s - s \;\; \& \;\; 1 \leq l_s + s < l_s + 2s \\ 2s - l_s & if \quad l_s - 2s \leq 0 < l_s - s \;\; \& \;\; l_s + s < l_s + 2s < 1 \\ 4s - 1 & if \quad l_s - 2s \leq 0 < l_s - s \;\; \& \;\; l_s + s < 1 \leq l_s + 2s \\ 2s + 1 - 2l_s & if \quad l_s - 2s \leq 0 < l_s - s \;\; \& \;\; 1 \leq l_s + s < l_s + 2s \\ l_s & if \quad l_s - 2s < l_s - s \leq 0 \;\; \& \;\; l_s + s < l_s + 2s < 1 \\ 2l_s + 2s - 1 & if \quad l_s - 2s < l_s - s \leq 0 \;\; \& \;\; l_s + s < 1 \leq l_s + 2s \\ 1 & if \quad l_s - 2s < l_s - s \leq 0 \;\; \& \;\; 1 \leq l_s + s < l_s + 2s \end{cases} \tag{3}
$$

Once we know the proportion of the population displaced and the fraction of the population that refugees represent, the fraction that remain internally displaced is computed as the difference.

This model delivers the fractions of internally and externally displaced populations as functions of the location and the size of the shock. However, in order to be able to use our model for empirical analysis, it is convenient to integrate out the location variable. With this purpose, we assume the shock location is equally likely to take place at any point in Country B. Appendix 4 shows that, integrating over the shock location, we obtain the conditional expectation of the fraction of people displaced, the fraction externally displaced and, by subtraction, the fraction internally displaced

$$
E_{l_s}[D|s] = s(2-s) \quad if \quad s \in [0,1] \tag{4}
$$

$$
E_{l_s}[R|s] = \begin{cases} 2s^2 & if \quad s \in [0, 0.5] \\ 4s - 2s^2 - 1 & if \quad s \in (0.5, 1] \end{cases} \tag{5}
$$

$$
E_{l_s}[I|s] = \begin{cases} s(2-3s) & if \quad s \in [0, 0.5] \\ s^2 - 2s + 1 & if \quad s \in (0.5, 1] \end{cases} \tag{6}
$$

Notice that, even though equations 2 and 3 are continuous but not differentiable functions, integrating over the location removes the non differentiability.

Figure 10 plots the expected values of the proportions of displaced people, refugees and IDPs as continuous and differentiable functions of the shock intensity. In its simplest form, the model delivers the following results. First, both the expected values of the proportions of displaced persons and refugees are increasing functions of the size of the shock. Second, the proportion of IDPs is increasing in the size of the shock for small values of it, has a maximum at $s = 1/3$ and from that point on it is decreasing in the size of the shock.

Finally, notice that these functions return fractions of population, thus the total number of people displaced, the total number of refugees and the total number of IDPs can be obtained multiplying these fractions by population. Therefore, population is in this model a natural measure of exposure.

2.1 Country size, ruggedness and proximity

The model developed in the previous section assumes the armed conflict shock is the only cause of displacement. In this section, we introduce various spatial variables into our analysis: country size, ruggedness and proximity. These variables are not causal forces of displacement but they modify the relationship between the shock and displacement and have some bearing in determining how total displacement splits into internal and external.

The area of a country plays an important role concerning internal and external displacement. Intuitively, holding everything else constant, displaced people should find it easier to escape from smaller countries. So far, the model has been developed under the assumption that the size of the Country B is unity. For simplicity, we assume that the largest country on earth has size unity and the other countries are a fraction a of the largest. Therefore, we now assume that location l_s ranges from zero to a, a value lower than one. Stretching the language, we will refer to a as the "area" of the country, despite the fact our model represents a country as a segment in the real line.

Next we make two assumptions. First, we assume population is uniformly distributed in the $[0,a]$ segment, so the probability density function of population is $1/a$ at each location in $[0,a]$ and zero otherwise. Notice the difference between the standard concept of "population density", say the number of people per square kilometer, and the probability density of the variable population, which describes how population is distributed within the country. We assume the latter is uniformly distributed but say nothing about the former, that is, the country could have a high or a low "population density". Second, without loss of generality, the size of the shock is now assumed to lay in the interval $[0,a]$. Notice that a shock of size a affects the entire population of the country, so it seems unnecessary to analyze larger shocks.

A geographical factor that affects internal and external displacement is the terrain ruggedness of the source country. Rugged source countries impose an extra difficulty for displacement. Let r denote the degree of ruggedness of a country, a value in the unit interval, with $r = 0$ for a perfectly flat country and $r = 1$ for an abrupt country highly unaccessible. We assume that in a rugged country, a shock of size s forces people affected to move a fraction γ of the distance displaced in the baseline model, that is, γs, where $\gamma = \gamma(r)$, with $\gamma'(r) < 0$, so the fraction γ is decreasing in ruggedness. In addition, we assume $\gamma(0) = 1$, so in a perfectly flat country, displaced persons move as described in the baseline model. We also assume that $0 < \gamma(1) < 1$, that is, in a country with the highest ruggedness parameter, $r = 1$, the fraction γ is small but positive.

Distance between countries is another geographical factor of relevance for the analysis of internal and external displacement. Intuitively, holding everything else constant, greater distance between two countries ought to be associated with lower refugee flows. However, the model analyzed in the previous section does not account for distance between countries. In fact, the model assumes countries A, B and C are contiguous. Let us introduce distance into the analysis. Let d_{ij} be the normalized distance between country i and country j, with $i, j = A, B, C$. Distance is normalized so that the longest distance between two countries equals unity. However, we continue using the contiguous countries design and model distance between countries as an iceberg-type cost. More specifically, out of all displaced people who reach the border between countries i and j, the fraction that crosses the border is a decreasing function of the distance between countries i and j, $\alpha_{ij} = \alpha\left(d_{ij}\right)$, with $\alpha'\left(d_{ij}\right) < 0$. This fraction equals one when distance is zero, $\alpha(0) = 1$, and is below one and positive when distance between countries i and j is maximal, $0 < \alpha(1) < 1$.

Modeling distance this way abstracts from whether countries are linked by land or necessarily by sea. In addition, it could be argued that this iceberg-type cost associated with crossing a border should be taken into account at the time of deciding whether to move or not. However, the model assumes that the decision whether to move or not is independent of the cost of crossing a border. The model also abstracts from asylum application costs and the destination country stance towards refugees. Importantly, notice that the fraction $1 - \alpha\left(d_{ij}\right)$ of those who reach the border between countries i and j pile up at the border, in line with a real world observation.

Under these assumptions, total displacement is

$$D = \begin{cases} 2\gamma s/a & if \quad l_s - \gamma s \geq 0 \quad \& \quad l_s + \gamma s \leq a \\ \left(l_s + \gamma s\right)/a & if \quad l_s - \gamma s < 0 \quad \& \quad l_s + \gamma s \leq a \\ 1 - \left(l_s - \gamma s\right)/a & if \quad l_s - \gamma s \geq 0 \quad \& \quad l_s + \gamma s > a \\ 1 & if \quad l_s - \gamma s < 0 \quad \& \quad l_s + \gamma s > a \end{cases} \tag{7}$$

Similarly, the fraction of country B's population that become refugees is

$$R = \begin{cases} 0 & if \quad 0 < l_s - 2\gamma s < l_s - \gamma s \quad \& \quad l_s + \gamma s < l_s + 2\gamma s < a \\ \alpha_{BC}(l_s + 2\gamma s - a)/a & if \quad 0 < l_s - 2\gamma s < l_s - \gamma s \quad \& \quad l_s + \gamma s < a \leq l_s + 2\gamma s \\ \alpha_{BC}(a - l_s)/a & if \quad 0 < l_s - 2\gamma s < l_s - \gamma s \quad \& \quad a \leq l_s + \gamma s < l_s + 2\gamma s \\ \alpha_{AB}(2\gamma s - l_s)/a & if \quad l_s - 2\gamma s < 0 \leq l_s - \gamma s \quad \& \quad l_s + \gamma s < l_s + 2\gamma s < a \\ [(2\gamma s - l_s)\alpha_{AB} + (l_s + 2\gamma s - a)\alpha_{BC}]/a & if \quad l_s - 2\gamma s < 0 \leq l_s - \gamma s \quad \& \quad l_s + \gamma s < a \leq l_s + 2\gamma s \\ [(2\gamma s - l_s)\alpha_{AB} + (a - l_s)\alpha_{BC}]/a & if \quad l_s - 2\gamma s < 0 \leq l_s - \gamma s \quad \& \quad a \leq l_s + \gamma s < l_s + 2\gamma s \\ \alpha_{AB} l_s/a & if \quad l_s - 2\gamma s < l_s - \gamma s \leq 0 \quad \& \quad l_s + \gamma s < l_s + 2\gamma s < a \\ [l_s\alpha_{AB} + (l_s + 2\gamma s - a)\alpha_{BC}]/a & if \quad l_s - 2\gamma s < l_s - \gamma s \leq 0 \quad \& \quad l_s + \gamma s < a \leq l_s + 2\gamma s \\ [l_s\alpha_{AB} + (a - l_s)\alpha_{BC}]/a & if \quad l_s - 2\gamma s < l_s - \gamma s \leq 0 \quad \& \quad a \leq l_s + \gamma s < l_s + 2\gamma s \end{cases} \tag{8}$$

Notice that total displacement depends on how many people are affected by the shock and it does not depend on whether they cross a border or not. Therefore, total displacement does not depend on distance. However, the fraction of country B's population that become refugees does depend on distance.

Appendix 4 shows how integrating out the location variable the expected values of total displacement, refugees and IDPs are

$$E_{l_s}[D|s,a,r,d_{AB},d_{BC}] = (2a\gamma s - \gamma^2 s^2)/a^2 \quad if \quad s \in [0,a]$$

$$E_{l_s}[R|s,a,r,d_{AB},d_{BC}] = \begin{cases} \gamma^2 s^2(\alpha_{AB} + \alpha_{BC})/a^2 & if \quad s \in [0, a/2] \\ (2a\gamma s - \gamma^2 s^2 - a^2/2)(\alpha_{AB} + \alpha_{BC})/a^2 & if \quad s \in (a/2, a] \end{cases}$$

$$E_{l_s}[I|s,a,r,d_{AB},d_{BC}] = \begin{cases} [2a\gamma s - \gamma^2 s^2 - \gamma^2 s^2 (\alpha_{AB} + \alpha_{BC})]/a^2 & if \quad s \in [0, a/2] \\ [2a\gamma s - \gamma^2 s^2 - (2a\gamma s - \gamma^2 s^2 - a^2/2)(\alpha_{AB} + \alpha_{BC})]/a^2 & if \quad s \in (a/2, a] \end{cases}$$

where the expected values of the relevant fractions are conditional on the size of the shock, area, ruggedness and distances to neighboring countries. Notice that the term $P_B = \alpha_{AB} + \alpha_{BC}$ can be interpreted as a measure of proximity.

Appendix 4 shows the derivatives of the expected value of the proportions of refugees, internally displaced and total displacement with respect to area, ruggedness and proximity. The signs of these derivatives indicate that the fraction of the population displaced responds negatively to the area and ruggedness while it is unrelated to proximity. The proportion of refugees also responds negatively to area and ruggedness, but it is positively related with proximity. However, the signs of the derivatives of the proportion of IDPs with respect to area, ruggedness and proximity depend on the magnitude of the size of the shock, area, proximity and ruggedness in a complex manner.

3 The data, empirical methods and evidence

3.1 The econometric methods

The dependent variables analyzed are total displacement, total number of refugees and IDPs. We use lagged population as exposure, therefore the dependent variable is the fraction of the population that refugees (respectively, IDPs or total displacement) represent. The expected value of the fraction of a country's population that become refugees (respectively, IDPs or displaced) is assumed to be an exponential function of a vector of covariates. The exponential regression functions are estimated by the Generalized Method of Moments (GMM), which is numerically identical to the pooled Poisson maximum likelihood estimator. We will term this estimator as Pooled-GMM. However, the Pooled-GMM estimator does not account for unobserved heterogeneity. Accounting for unobserved time-invariant heterogeneity via country of origin fixed effects cannot simultaneously identify the effect of magnitudes such as area or terrain roughness which are time invariant. Thus, even though our empirical evidence is based on a panel data set, we cannot resort to using the popular fixed effects methodology, which would preclude inference about time invariant covariates, such as area, ruggedness and proximity. Instead, we use the Pre-Sample Mean Generalized Method of Moments (PSM-GMM) suggested by Blundell et al. (2002). The PSM-GMM estimation method is like the Pooled-GMM only that it includes an additional regressor: the average value of the dependent variable during a period prior to the sample period. The pre-sample mean of the dependent variable is used as a proxy for the unobserved heterogeneity. As the dependent variable is measured as a fraction of the population, the pre-sample mean is also measured as a fraction of the population.

3.2 The data

Variable definitions and data sources are shown in Appendix 4. In this paper we use refugee and IDP populations from the UNHCR populations of concern database. Figure 11 shows world aggregate time series of the number of IDPs and refugees. Over the last few decades, the number of IDPs has increased remarkably, while the number of refugees has stabilized. The total number of forcibly displaced people keeps on growing and represents one of the most important problems for humanity. Forced displacement is not a problem of just poor and conflicted countries, it affects almost every country on earth. Figure 12 plots the number of countries listed by the UNHCR as the origin of some refugees. This number experienced a fairly constant growth until the mid 1980s, grew faster during the late 1980s and 1990s and reached almost all the countries by the early 2000's. The empirical evidence reported below refers to the 1995-2016 period, including 200 countries and territories listed in Table 1. Displacement counts prior to 1995 are used to compute the pre-sample mean.

The data includes only IDPs uprooted by conflict and human rights violations. Thus, the data does not include displacement caused by natural disasters.[2] Refugee data are dyadic time series, i.e. each refugee count is associated with an origin country, a destination country and a year. IDP counts are longitudinal data, each involving a country and year. We dropped all refugee counts whose origin were unknown and replaced missing refugee or IDP counts by zero counts.[3] Next, we aggregated the refugee counts adding all destinations to obtain an unbalanced longitudinal origins dataset.

In addition to IDPs and refugee counts, our data set includes the following covariates. Population data from the World Bank, which is used as a measure of exposure, that is, we measure the number of refugees, IDPs and total displacement as a fraction of the origin country's population. Area and the average Riley et al. (1999) index of terrain ruggedness are from Nunn and Puga (2012). Distances between each country's most populated cities are from CEPII. For the empirical analysis, a country's proximity, P_i, is measured as an exponential average of distances from origin country i to all the other N destination countries

$$P_i = \sum_{j=1}^{N} e^{-d_{ij}}$$

where d_{ij} represents the distance between country i and country j, measured as a fraction of the largest distance.[4] Conflict intensity data, from the UCDP-PRIO conflict dataset, are measured as a categorical variable with three categories corresponding to no conflict, low intensity conflict and high intensity conflict.

In addition to these covariates, we also include other control variables regarding which the model does not contribute anything, but have previously been included in empirical models of displacement. First, a large share of forced displacement does not have conflict as the root cause, instead the cause forcing migration is the lack of civil liberties or political rights. Therefore, we include categorical measures of civil liberties and political rights from Freedom House. Second, it is often argued that refugee counts hide economic migrants, therefore, refugee counts should be negatively related with economic conditions in the source country. Thus, we include GDP per capita from the World Bank in the empirical specification. Third, another forced migration correlate reported in the literature is colonial relationship. Thus, we include the number of countries with which a source country had a colonial relationship (either as a colony or colonizer). Fourth, in the same vein, we also control for the number of countries with which a source country shares language/ethnicity group in common with at least 9% of the population.

The empirical evidence is complemented by including in the analysis other spatial covariates that we did not account for in the theoretical model but might have a bearing in forced displacement determination. A potentially relevant spatial variable is whether a country's area includes a large desert. Our model assumes population is uniformly distributed in a country's territory. However, the presence of a desert invalidates our assumption. To account for this issue, we include an indicator of whether more than 20% of a country's area is a desert. Another spatial feature that is not incorporated in our model is whether a country is landlocked or not. Arguably, a landlocked country, surrounded by other countries, must be easier to scape from than an island country. Therefore, we also include in the analysis an indicator of whether a country is landlocked or not.

3.3 Empirics

In this section we report empirical evidence using data from 1995 to 2016. Table 2 reports exponential regression estimates for refugee counts. Column (1) reports Pooled-GMM estimates and includes only the spatial variables and conflict intensity, the only variables accounted for in the theoretical model. As predicted by the model, area has a negative and significant effect and proximity a positive and significant one. Ruggedness has a positive and not statistically significant coefficient, while the model predicts a negative partial effect. In accordance with the predictions of the model, conflict intensity has a positive effect on external displacement: high intensity conflicts have a positive and significant effect on refugee counts, while low intensity ones do not affect refugee counts significantly.

Next we focus on the PSM-GMM estimates reported in columns (2), (3) and (4). These exponential regressions include the (log of) the mean value of the endogenous variable (the number of refugees as a fraction of population) during the pre-sample period, in our case, 1960-1994. Therefore, the sample is restricted to 168 countries with positive pre-sample mean, so that the logarithm can be computed. Adding the pre-sample mean to account for unobserved heterogeneity in column (2) improves the goodness of fit considerably, as measured by the pseudo-R^2. In addition, the coefficient on ruggedness turns out to be negative, as predicted by the model, although it is significant only at ten percent. Furthermore, the low intensity conflict indicator also becomes significant at the ten per cent level. Finally, the pre-sample mean is positive and significant, thus indicating its relevance in accounting for heterogeneity.

Column (3) in Table 2 includes other determinants of refugee counts found relevant in the literature. Civil liberties are measured on a scale of 1 to 7, with 1 representing the highest standards of civil liberties and 7 the lowest. We incorporate this variable using 6 dummy variables for levels 2 to 7, leaving level 1 (the highest standards) as the reference group. It turns out that as we move to lower civil liberties standards (from level 2 to level 7) the estimated coefficients are larger. In other words, refugee counts are uniformly increasing as civil liberties standards deteriorate. On the other hand, with regard to the effect of economic conditions at the source countries, we find that higher levels of real GDP per capita in origin countries are significantly associated with lower refugee counts. However, neither the common linguistic-ethnicity proxy nor the colonial-relation measure turn out to be significant. Overall, including civil liberties and GDP per capita improves the goodness of fit and increases the significance of the spatial and conflict intensity variables.

Column (4) accounts for other spatial variables not considered in the theoretical model that might be relevant for refugee flows in practice. Neither the desert nor the landlocked indicators are significant, while the results for the other covariates remain almost invariant quantitatively and in terms of significance.

Replacing the civil liberties with political rights yields very similar results, which are not reported and available as an online appendix.

Table 3 reports the same exponential regressions as in Table 2 with two differences. First, the dependent variable is now IDP counts. Second, Data on IDPs are available from 1993 to 2016. Thus, the pre-sample period includes only IDP counts for 1993 and 1994, which leaves just two IDP counts to compute the pre-sample mean. However, under the assumptions of our model, at the early stages of displacement, all people displaced are IDPs and at later stages some become refugees. A positive refugee count is therefore evidence of earlier internal displacement. This reasoning suggests that refugee counts must be preceded by IDP counts. Accordingly, the pre-sample mean used in columns (2) to (4) corresponds to the pre-sample mean of total displacement, that is, the sum of refugees and IDPs.

The results reported in Table 3 are very similar to those reported in Table 2, both quantitatively and in terms of significance. The only differences are that civil liberties and GDP per capita are somewhat less significant for IDP regressions than for refugee ones. For instance, civil liberties level 2 (a level below the highest standards of civil liberties) is not significantly different from the reference group (civil liberties level 1, the highest standards of civil liberties). However, civil liberties level 7 (the lowest standards) is highly significant and has a quantitatively similar coefficient estimate to the refugees regression. IDP counts are not strictly uniformly decreasing in civil liberties, but almost.

In accordance with the model's prediction, the coefficients on the conflict intensity indicators are positive and statistically significant. Coefficient estimates on area, ruggedness and proximity are significant and have the same signs as in the refugee regression. As shown in appendix 4, the model predicts that the partial derivatives of IDPs with respect to the spatial variables might be either positive or negative depending on the values of all the spatial variables and conflict intensity. According to the evidence reported in Table 3, area has a negative partial effect, which is consistent with the theory provided proximity is not very high. Ruggedness affects internal displacement negatively and significantly, which is in line with the model only for low values of proximity. Proximity has a positive coefficient on the IDPs regression, as predicted by the model for high shock intensity. While this evidence is compatible with the predictions of the model, strictly speaking, we cannot validate nor refute the model based on the results from the IDPs regressions.

Table 4 reports exponential regression results for total displacement which are similar to those found for internal and external displacement. In particular, the spatial variables, area, ruggedness and proximity are significant and have the same signs as in the previous tables. Area and ruggedness have a negative effect on total displacement, as suggested by the model. However, the model predicts proximity should not have any bearing in determining total displacement, while we find a positive and significant effect. A feasible explanation of this outcome is that total displacement and, in particular, refugee counts record not only conflict induced displacement, as assumed in the model, but also those who flee their country for political, religious, ethnic, gender or other reasons, and also might include some economic migrants. All these people might be more likely to flee their country of origin for higher proximity levels.

Results not reported but available upon request show that the findings reported in Tables 2-4 remain very much the same when the civil liberties indicator is replaced with a political rights indicator.

Overall, our reading of the results is that they fundamentally agree with the predictions of the theoretical model. First, all the evidence uses population as exposure, as indicated by the model, and the specification choice results in good fit for this type of data. Second, the spatial variables have the predicted coefficient signs in the regressions, with the only exception of proximity in the total displacement regressions. Third, all displacement counts (refugees, IDPs and total) are uniformly increasing in conflict intensity. Finally, the other determinants of forced migration included in the regressions but not accounted for in the theoretical model have coefficient signs which accord with previous findings in the literature.

4 Conclusions and directions for further research

This article proposes a spatial model to analyze internal and external forced migration, a topic of increasing relevance nowadays. The bare bones model worked out in this article does not pretend to develop a new theory about how internal and external forced displacement are generated. On the contrary, it recognizes that previous studies have shown that armed conflict and lack of political rights or civil liberties cause people displacement. This is to say that the root causes of forced displacement are well known. The contribution of this paper is to analyze the role of spatial factors in the relationship between the root causes (armed conflict and lack of liberties) and forced displacement. In the causal inference jargon, these spatial factors are called effect modifiers.

The model predicts how some spatial factors are related with internal, external and total displacement. In particular, we focus on country size (measured by area), how abrupt the geography of a country is (measured by its ruggedness) and the degree of geographical proximity of the country with respect to the rest of the world. The model also predicts how conflict intensity ought to be related to internal displacement and forced migration. In addition, the model suggests that population should be used as a measure of exposure, and that is how it enters the empirical analysis. We test these predictions against real data. The evidence reported is broadly consistent with the predictions of the model.

We chose to develop a barebones model intentionally, hoping that its simplicity would seduce other scholars. There are a number of feasible extensions to this work. First, the model can be developed using other spatial frameworks by assuming that the origin country is, say a circle, or other two- or higher-dimensional figures. The model could also be solved under different assumptions about how shocks affect people, induced displacement, costs of displacement, cost of settlement and other issues we have not taken into account in our simplified model. The model could also be used to model displacement provoked by natural disasters and the predictions of such a model could be tested using data from displacement induced by natural disasters. Finally, although our model focuses on the country of origin, the analysis could be extended to destination countries, or even to dyadic flows.

References

Akgündüz, Y. E. and Torun, H. (2018). Two and a half million syrian refugees, skill mix and capital intensity. GLO Discussion Paper Series 186, Global Labor Organization (GLO).

Alix-Garcia, J., Bartlett, A., and Saah, D. (2012). Displaced populations, humanitarian assistance and hosts: A framework for analyzing impacts on semi-urban households. *World Development*, 40(2):373 – 386.

Alix-Garcia, J., Bartlett, A., and Saah, D. (2013). The landscape of conflict: IDPs, aid and land-use change in Darfur. *Journal of Economic Geography*, 13(4):589–617.

Alix-Garcia, J. and Saah, D. (2009). The effect of refugee inflows on host communities: Evidence from tanzania. *World Bank Economic Review*, 24:148–170.

Ayata, B. and Yukseker, D. (2005). A belated awakening: national and international responses to the internal displacement of kurds in turkey. *New perspectives on Turkey*, 32:5–42.

Baez, J. E. (2011). Civil wars beyond their borders: The human capital and health consequences of hosting refugees. *Journal of Development Economics*, 96(2):391 – 408.

Ball, P., Betts, W., Scheuren, F., Dudukovich, J., and Asher, J. (2002). Political killings in kosova/kosovo, march-june 1999. Technical report, American Association for the Advancement of Science.

Barthel, F. and Neumayer, E. (2015). Spatial dependence in asylum migration. *Journal of Ethnic and Migration Studies*, 41(7):1131–1151.

Blundell, R., Griffith, R., and Windmeijer, F. (2002). Individual effects and dynamics in count data models. *Journal of Econometrics*, 108(1):113 – 131.

Böhmelt, T., Bove, V., and Gleditsch, K. S. (2019). Blame the victims? refugees, state capacity, and non-state actor violence. *Journal of Peace Research*, 56(1):73–87.

Bohnet, H., Cottier, F., and Hug, S. (2018). Conflict-induced idps and the spread of conflict. *Journal of Conflict Resolution*, 62(4):691–716.

Borjas, G. J. (2017). The wage impact of the marielitos: A reappraisal. *ILR Review*, 70(5):1077–1110.

Card, D. (1990). The impact of the mariel boatlift on the miami labor market. *ILR Review*, 43(2):245–257.

Ceritoglu, E., Yunculer, H. B. G., Torun, H., and Tumen, S. (2017). The impact of syrian refugees on natives' labor market outcomes in turkey: evidence from a quasi-experimental design. *IZA Journal of Labor Policy*, 6(1):5.

Crisp, J. (2010). Forced displacement in africa: Dimensions, difficulties, and policy directions. *Refugee Survey Quarterly*, 29(3):1–27.

Czaika, M. (2009). *The Political Economy of Refugee Migration and Foreign Aid*. Palgrave Macmillan.

Czaika, M. and Kis-Katos, K. (2009). Civil conflict and displacement: Village-level determinants of forced migration in aceh. *Journal of Peace Research*, 46(3):399–418.

Davenport, C., Moore, W., and Poe, S. (2003). Sometimes you just have to leave: Domestic threats and forced migration, 1964-1989. *International Interactions*, 29(1):27–55.

Depetris-Chauvin, E. and Santos, R. J. (2018). Unexpected guests: The impact of internal displacement inflows on rental prices in colombian host cities. *Journal of Development Economics*, 134:289 – 309.

Dreher, A., Fuchs, A., and Langlotz, S. (2019). The effects of foreign aid on refugee flows. *European Economic Review*, 112:127 – 147.

Echevarria, J. and Gardeazabal, J. (2016). Refugee gravitation. *Public Choice*, 169(3-4):269–292.

Engel, S. and Ibáñez, A. M. (2007). Displacement due to violence in colombia: A household level analysis. *Economic Development and Cultural Change*, 55(2):335–365.

Esen, O. and Oğuş Binatlı, A. (2017). The impact of syrian refugees on the turkish economy: Regional labour market effects. *Social Sciences*, 6(4).

Fernandes, W. (2017). Internally displaced persons and northeast india. *International Studies*, 50(4):287–305.

Fisk, K. (2019). Camp settlement and communal conflict in sub-saharan africa. *Journal of Peace Research*, 56(1):58–72.

Goswami, U. (2007). Internal displacement migration and policy in northeastern india.

Havinga, T. and Böcker, A. (1999). Country of asylum by choice or by chance: Asylum-seekers in belgium, the netherlands and the uk. *Journal of Ethnic and Migration Studies*, 25(1):43–61.

Hotelling, H. (1929). Stability in competition. *The Economic Journal*, 39(153):41–57.

Ibáñez, A. M. and Vélez, C. E. (2008). Civil conflict and forced migration: The micro determinants and welfare losses of displacement in colombia. *World Development*, 36(4):659 – 676.

Iqbal, Z. (2007). The geo-politics of forced migration in africa, 1992–2001. *Conflict Management and Peace Science*, 24(2):105–119.

Kälin, W. (2008). Guiding principles on internal displacement. *Stud. Transnat'l Legal Pol'y*, 38:1.

Kondylis, F. (2008). Agricultural outputs and conflict displacement: Evidence from a policy intervention in rwanda. *Economic Development and Cultural Change*, 57(1):31–66.

Kondylis, F. (2010). Conflict displacement and labor market outcomes in post-war bosnia and herzegovina. *Journal of Development Economics*, 93(2):235 – 248.

Lanjouw, S., Mortimer, G., and Bamforth, V. (2000). Internal displacement in burma. *Disasters*, 24(3):228–239.

Lischer, S. K. (2008). Security and displacement in iraq: responding to the forced migration crisis. *International Security*, 33(2):95–119.

McEniry, M., Samper-Ternent, R., and Cano-Gutierrez, C. (2019). Displacement due to armed conflict and violence in childhood and adulthood and its effects on older adult health: The case of the middle-income country of colombia. *SSM - Population Health*, 7:100369.

Milton, D., Spencer, M., and Findley, M. (2013). Radicalism of the hopeless: Refugee flows and transnational terrorism. *International Interactions*, 39(5):621–645.

Moore, W. H. and Shellman, S. M. (2007). Whither will they go? a global study of refugees' destinations, 1965–1995. *International Studies Quarterly*, 51(4):811–834.

Morales, J. S. (2018). The impact of internal displacement on destination communities: Evidence from the colombian conflict. *Journal of Development Economics*, 131:132 – 150.

Muggah, R. (2003). A tale of two solitudes: Comparing conflict and development-induced internal displacement and involuntary resettlement. *International Migration*, 41(5):5–31.

Munive, J. (2019). Resilience in Displacement and the Protection of Civilians in South Sudan. *Journal of Refugee Studies*, pages 1–21.

Neumayer, E. (2005a). Asylum recognition rates in western europe: Their determinants, variation, and lack of convergence. *Journal of Conflict Resolution*, 49(1):43–66.

Neumayer, E. (2005b). Bogus refugees? the determinants of asylum migration to western europe. *International Studies Quarterly*, 49(3):389–410.

Nunn, N. and Puga, D. (2012). Ruggedness: The blessing of bad geography in africa. *Review of Economics and Statistics*, 94(1):20–36.

Peri, G. and Yasenov, V. (2019). The labor market effects of a refugee wave:synthetic control method meets the mariel boatlift. *Journal of Human Resources*.

Riley, S., Degloria, S., and Elliot, S. (1999). A terrain ruggedness index that quantifies topographic heterogeneity. *Internation Journal of Science*, 5:23–27.

Roper, S. D. and Barria, L. A. (2010). Burden sharing in the funding of the unhcr: Refugee protection as an impure public good. *Journal of Conflict Resolution*, 54(4):616–637.

Rüegger, S. (2019). Refugees, ethnic power relations, and civil conflict in the country of asylum. *Journal of Peace Research*, 56(1):42–57.

Salehyan, I. (2008). The externalities of civil strife: Refugees as a source of international conflict. *American Journal of Political Science*, 52(4):787–801.

Salehyan, I. and Gleditsch, K. S. (2006). Refugees and the spread of civil war. *International Organization*, 60(02):335–366.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of mathematical sociology*, 1(2):143–186.

Tumen, S. (2016). The economic impact of syrian refugees on host countries: Quasi-experimental evidence from turkey. *American Economic Review*, 106(5):456–60.

UNHCR (2006). Oau convention governing the specific aspects of refugee problems in africa.

UNHCR (2011). United nations high comisioner for the refugees: Convention and protocol relating to the status of refugees.

Uzonyi, G. (2015). Refugee flows and state contributions to post-cold war un peacekeeping missions. *Journal of Peace Research*, 52(6):743–757.

Verme, Paolome, P. and Schuettler, K. (2019). The impact of forced displacement on host communities : A review of the empirical literature in economics. Technical report, World Bank.

Appendixes

Mathematical appendix

To integrate out the location we proceed in two regions defined by the intersection of the two lines in Figure 2. First we integrate 2 for values of the shock in the interval $[0,\frac{1}{2}]$ and then in the interval $\left(\frac{1}{2},1\right]$

$$E_{l_s}[D|s] = \begin{cases} \int_s^{1-s} 2s\mathrm{d}l_s + \int_0^s (l_s+s)\mathrm{d}l_s + \int_{1-s}^1 (s+1-l_s)\mathrm{d}l_s = 2s-s^2 & if \quad s\in[0,\frac{1}{2}] \\ \int_0^{1-s} 2s\mathrm{d}l_s + \int_s^1 (s+1-l_s)\mathrm{d}l_s + \int_{1-s}^s dl_s = 2s-s^2 & if \quad s\in\left(\frac{1}{2},1\right] \end{cases}$$

Next we integrate 3 in four regions defined by the intersections of the four lines in Figure 6. For values of the shock in the intervals $\left[0,\frac{1}{4}\right]$, $\left(\frac{1}{4},\frac{1}{3}\right]$, $\left(\frac{1}{3},\frac{1}{2}\right]$ and $\left(\frac{1}{2},1\right]$

$$E_{l_s}[R|s] = \begin{cases} \int_0^s (l_s)\mathrm{d}l_s + \int_s^{2s} (2s-l_s)\mathrm{d}l_s \\ \quad + \int_{1-2s}^{1-s} (l_s+2s-1)\mathrm{d}l_s + \int_{1-s}^1 (1-l_s)\mathrm{d}l_s & if \quad s\in[0,\frac{1}{4}] \\ \int_0^s (l_s)\mathrm{d}l_s + \int_s^{1-2s} (2s-l_s)\mathrm{d}l_s + \int_{1-2s}^{2s} (4s-1)\mathrm{d}l_s \\ \quad + \int_{2s}^{1-s} (l_s+2s-1)\mathrm{d}l_s + \int_{1-s}^1 (1-l_s)\mathrm{d}l_s & if \quad s\in(\frac{1}{4},\frac{1}{3}] \\ \int_0^{1-2s} (l_s)\mathrm{d}l_s + \int_{1-2s}^s (2s+2l_s-1)\mathrm{d}l_s + \int_s^{1-s} (4s-1)\mathrm{d}l_s \\ \quad + \int_{1-s}^{2s} (2s+1-2l_s)\mathrm{d}l_s + \int_{2s}^1 (1-l_s)\mathrm{d}l_s & if \quad s\in(\frac{1}{3},\frac{1}{2}] \\ \int_0^{1-s} (2l_s+2s-1)\mathrm{d}l_s + \int_{1-s}^s dl_s + \int_s^1 (2s+1-2l_s)\mathrm{d}l_s & if \quad s\in(\frac{1}{2},1] \end{cases}$$

which results in the much simpler expression

$$E_{l_s}[R|s] = \begin{cases} 2s^2 & if \quad s\in[0,\frac{1}{2}] \\ 4s-2s^2-1 & if \quad s\in(\frac{1}{2},1] \end{cases}$$

The fraction of the population that remain as IDPs is found by subtraction

$$E_{l_s}[I|s] = E[D]-E[R] = \begin{cases} 2s-3s^2 & if \quad s\in[0,\frac{1}{2}] \\ s^2-2s+1 & if \quad s\in(\frac{1}{2},1] \end{cases}$$

Next consider the model of Section 2.1. Integrating equation 7

$$E_{l_s}[D|s,a,r,d_{AB},d_{BC}] =$$

$$= \begin{cases} [\int_0^a (2\gamma s)/a\,\mathrm{d}l_s - \int_0^{\gamma s}(\gamma s - l_s)/a\,\mathrm{d}l_s - \int_{a-\gamma s}^{a}(l_s+\gamma s-a)/a\,\mathrm{d}l_s]/a & if \quad s \in [0,a/2] \\ [\int_0^a (2\gamma s)/a\,dl_s - \int_0^{a-\gamma s}(\gamma s - l_s)/a\,\mathrm{d}l_s - \int_{\gamma s}^{a}(l_s+\gamma s-a)/a\,dl_s \\ \quad - \int_{a-\gamma s}^{\gamma s}[(\gamma s - l_s)+(l_s+\gamma s-a)]/a\,\mathrm{d}l_s]/a & if \quad s \in (a/2,a] \end{cases}$$

$$= (2\gamma a s - \gamma^2 s^2)/a^2 \quad \in [0,a]$$

Similarly integrating 8 we obtain

$$E_{l_s}[R|s,a,r,d_{AB},d_{BC}] =$$

$$= \begin{cases} [\int_0^{\gamma s}\alpha_{A,B}(l_s)/a\,\mathrm{d}l_s + \int_{\gamma s}^{2\gamma s}\alpha_{A,B}(2\gamma s - l_s)/a\,\mathrm{d}l_s \\ + \int_{a-2\gamma s}^{a-\gamma s}\alpha_{B,C}(l_s+2\gamma s-a)/a\,\mathrm{d}l_s + \int_{a-\gamma s}^{a}\alpha_{B,C}(a-l_s)/a\,\mathrm{d}l_s]/a & if \quad s \in [0,a/4] \\ [\int_0^{\gamma s}\alpha_{A,B}(l_s)/a\,\mathrm{d}l_s + \int_{\gamma s}^{a-2\gamma s}\alpha_{A,B}(2\gamma s - l_s)/a\,\mathrm{d}l_s \\ + \int_{a-2\gamma s}^{2\gamma s}[\alpha_{A,B}(2\gamma s - l_s) + \alpha_{B,C}(l_s+2\gamma s-a)]/a\,\mathrm{d}l_s \\ + \int_{2\gamma s}^{a-\gamma s}\alpha_{B,C}(l_s+2\gamma s-a)/a\,\mathrm{d}l_s + \int_{a-\gamma s}^{a}\alpha_{B,C}(a-l_s)/a\,dl_s]/a & if \quad s \in (a/4,a/3] \\ [\int_0^{a-2\gamma s}\alpha_{A,B}(l_s)/a\,\mathrm{d}l_s + \int_{a-2\gamma s}^{\gamma s}\alpha_{A,B}(l_s)/a + \alpha_{B,C}(l_s+2\gamma s-a)/a\,\mathrm{d}l_s \\ + \int_{\gamma s}^{a-\gamma s}\alpha_{A,B}(2\gamma s - l_s)/a + \alpha_{B,C}(l_s+2\gamma s-a)/a\,\mathrm{d}l_s \\ + \int_{a-\gamma s}^{2\gamma s}\alpha_{B,C}(2\gamma s + a - 2l_s)/a\,\mathrm{d}l_s + \int_{2\gamma s}^{a} 1/\alpha_{B,C}(a-l_s)/a\,\mathrm{d}l_s]/a & if \quad s \in (a/3,a/2] \\ + \int_{a-\gamma s}^{\gamma s}\alpha_{A,B}\,l_s/a[\int_0^{a-\gamma s}\alpha_{A,B}\,l_s/a + \alpha_{B,C}(l_s+2\gamma s-a)/a\,\mathrm{d}l_s \\ +\alpha_{B,C}(a-l_s)/a\,dl_s + \int_{\gamma s}^{a}\alpha_{A,B}(2\gamma s - l_s)/a + \alpha_{\mathrm{B,C}}(\mathrm{a}-\mathrm{l_s})/a\,\mathrm{d}l_s]/a \quad = & if \quad s \in (a/2,a] \end{cases}$$

$$= \begin{cases} [\gamma^2 s^2(\alpha_{A,B}+\alpha_{B,C})]/a^2 & if \quad s \in [0,a/4] \\ [\gamma^2 s^2(\alpha_{A,B}+\alpha_{B,C})]/a^2 & if \quad s \in (a/4,a/3] \\ [\gamma^2 s^2(\alpha_{A,B}+\alpha_{B,C})]/a^2 & if \quad s \in (a/3,a/2] \\ [(2\gamma a s - \gamma^2 s^2 - a^2/2)(\alpha_{A,B}+\alpha_{B,C})]/a^2 & if \quad s \in (a/2,a] \end{cases}$$

In particular, we are interested in the derivatives of (4), (5) and (6) with respect to area, ruggedness and proximity. These are as follows:

$$\frac{\partial E_{l_s}[D|s,a,\gamma(r),P_B]}{\partial a} = \frac{2\gamma s}{a^3}(\gamma s - a) \leq 0,$$

$$\frac{\partial E_{l_s}[D|s,a,\gamma(r),P_B]}{\partial r} = \frac{2\gamma s}{a^2}(a-\gamma s)\gamma'(r) \leq 0,$$

$$\frac{\partial E_{l_s}[D|s,a,\gamma(r),P_B]}{\partial P_B} = 0.$$

Ruggedness and country size reduce the fraction of the population displaced while proximity does not affect total displacement. The fraction of the population that become refugees depends on these magnitudes in a way that depends on whether the shock is below or above $a/2$. In particular we have that the derivatives of the fraction of the population that become refugees are

$$\frac{\partial E_{l_s}[R|s,a,\gamma(r),P_B]}{\partial a} = \begin{cases} -\frac{2\gamma^2 s^2 P_B}{a^3} \leq 0 & if \quad s \in [0,\frac{a}{2}] \\ \frac{2\gamma s(\gamma s - a)P_B}{a^3} < 0 & if \quad s \in (\frac{a}{2},a] \end{cases}$$

$$\frac{\partial E_{l_s}[R|s,a,\gamma(r),P_B]}{\partial r} = \begin{cases} \frac{2\gamma'(r)\gamma s P_B}{a^2} \leq 0 & if \quad s \in \left[0,\frac{a}{2}\right] \\ \frac{2s\gamma'(r)(a-\gamma s)P_B}{a^2} \leq 0 & if \quad s \in \left(\frac{a}{2},a\right] \end{cases}$$

$$\frac{\partial E_{l_s}[R|s,a,\gamma(r),P_B]}{\partial P_B} = \begin{cases} \left(\frac{\gamma s}{a}\right)^2 \geq 0 & if \quad s \in [0,\frac{a}{2}] \\ \frac{1}{2} - \frac{(\gamma s - a)^2}{a^2} > 0 & if \quad s \in (\frac{a}{2},a] \end{cases}$$

Therefore, country size and ruggedness affect negatively, and proximity positively, the fraction of the population that become refugees.

Finally, the fraction of the population that become IDPs depends on country size, ruggedness and proximity as follows

$$\frac{\partial E_{l_s}[I|s,a,\gamma(r),P_B]}{\partial a} = \begin{cases} \frac{2\gamma s(\gamma s(1+P_B)-a)}{a^3} & if \quad s \in [0,\frac{a}{2}] \\ \frac{2\gamma s(\gamma s - a)(1-P_B)}{a^3} & if \quad s \in (\frac{a}{2},a], \end{cases}$$

$$\frac{\partial E_{l_s}[I\,|s,a,\gamma(r),P_B]}{\partial r}=\begin{cases} -\frac{2\gamma'(r)s(\gamma s(1+P_B)-a)}{a^2} & if \quad s\in(0,\frac{a}{2}] \\ -\frac{2\gamma'(r)s(1-P_B)(\gamma s-a)}{a^2} & if \quad s\in(\frac{a}{2},a], \end{cases}$$

$$=\frac{\partial E_{l_s}[I\,|s,a,\gamma(r),P_B]}{\partial P_B}=\begin{cases} -\left(\frac{\gamma s}{a}\right)^2 & if \quad s\in(0,\frac{a}{2}] \\ \frac{1}{2}-\frac{\gamma s(2a-\gamma s)}{a^2} & if \quad s\in(\frac{a}{2},a]. \end{cases}$$

While the signs of the derivatives of total displacement and refugees are uniform on s, this is no longer the case for the signs of these derivatives of internal displacement

	$0\le P_B\le 1$	$1<P_B\le 2$	
		$0<\gamma\le\frac{2}{3}$	$\frac{2}{3}<\gamma\le 1$
$0\le s\le\frac{a}{2}$	$\frac{\partial E_{l_s}[I\,|s,a,\gamma(r),P_B]}{\partial a}\le 0$	$\frac{\partial E_{l_s}[I\,|s,a,\gamma(r),P_B]}{\partial a}\le 0$	$\frac{\partial E_{l_s}[I\,|s,a,\gamma(r),P_B]}{\partial a}>0$
$\frac{a}{2}<s\le a$	$\frac{\partial E_{l_s}[I\,|s,a,\gamma(r),P_B]}{\partial a}\le 0$	$\frac{\partial E_{l_s}[I\,|s,a,\gamma(r),P_B]}{\partial a}>0$	

	$0\le P_B\le 1$	$1<P_B\le 2$	
		$0<\gamma\le\frac{2}{3}$	$\frac{2}{3}<\gamma\le 1$
$0\le s\le\frac{a}{2}$	$\frac{\partial E_{l_s}[I\,|s,a,\gamma(r),P_B]}{\partial r}\le 0$	$\frac{\partial E_{l_s}[I\,|s,a,\gamma(r),P_B]}{\partial r}\le 0$	$\frac{\partial E_{l_s}[I\,|s,a,\gamma(r),P_B]}{\partial r}>0$
$\frac{a}{2}<s\le a$	$\frac{\partial E_{l_s}[I\,|s,a,\gamma(r),P_B]}{\partial r}\le 0$	$\frac{\partial E_{l_s}[I\,|s,a,\gamma(r),P_B]}{\partial r}>0$	

	$0 < \gamma < \frac{(2-\sqrt{2})a}{2s}$	$\frac{(2-\sqrt{2})a}{2s} \leq \gamma \leq 1$
$0 \leq s \leq \frac{a}{2}$	$\frac{\partial E_{l_s}[I \mid s,a,\gamma(r),P_B]}{\partial P_B} \leq 0$	
$\frac{a}{2} < s \leq a$	$\frac{\partial E_{l_s}[I \mid s,a,\gamma(r),P_B]}{\partial P_B} > 0$	$\frac{\partial E_{l_s}[I \mid s,a,\gamma(r),P_B]}{\partial P_B} \leq 0$

Data sources

The data set compiled is a blend of the following data sets:

1. Persons of concern Dataset. United Nations High Commissioner for the Refugees (UNHCR). Time frame: 1951-2016. Variables used: Refugees, IDPs.
   Availability: `http://popstats.unhcr.org/en/persons_of_concern`

2. Armed Conflict Dataset. Uppsala Conflict Data Program (UCDP) Uppsala University / Peace Research Institute Oslo (PRIO). Time frame: 1946-2015. Variables used: Location. The name(s) of the country/countries whose government(s) have a primary claim to the issue in dispute. Year of observation. The date when the conflict activity reached 25 battle-related deaths in a year. The date when conflict activity ended. The intensity level of the armed conflict.
   Availability: `https://www.prio.org/Data/Armed-Conflict/UCDP-PRIO/`

3. Freedom in the World Dataset. Freedom House. Time frame: 1972-2015. Variables used: Civil Liberties and Political Rights.
   Availability: `https://freedomhouse.org`

4. Geographical and Distance dataset. CEPII, SciencesPo Department of Economics. Dyadic data set. Variables used: Common language/ethnicity indicator. Colonial relationship indicator. Simple distance (most populated cities, km).
   Availability: `http://econ.sciences-po.fr/staff/thierry-mayer`

5. World Development Indicators. The World Bank. Population. GDP per capita, PPP (constant 2011 international $).
   Availability: `http://databank.worldbank.org/data/home.aspx`

6. Geographical variables. Land area. Riley's Index of Terrain Ruggedness. From Nunn and Puga (2012).
   Availability: `https://diegopuga.org/data/rugged/`

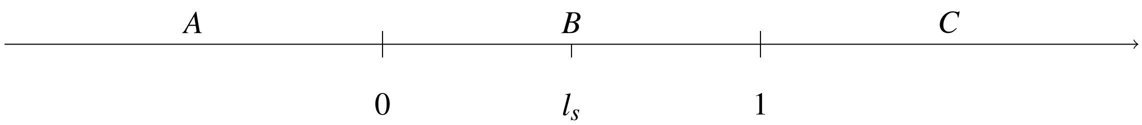

Figure 1: Countries A, B and C are located in the real line. Points 0 and 1 are the borders. The shock takes place at location l_s.

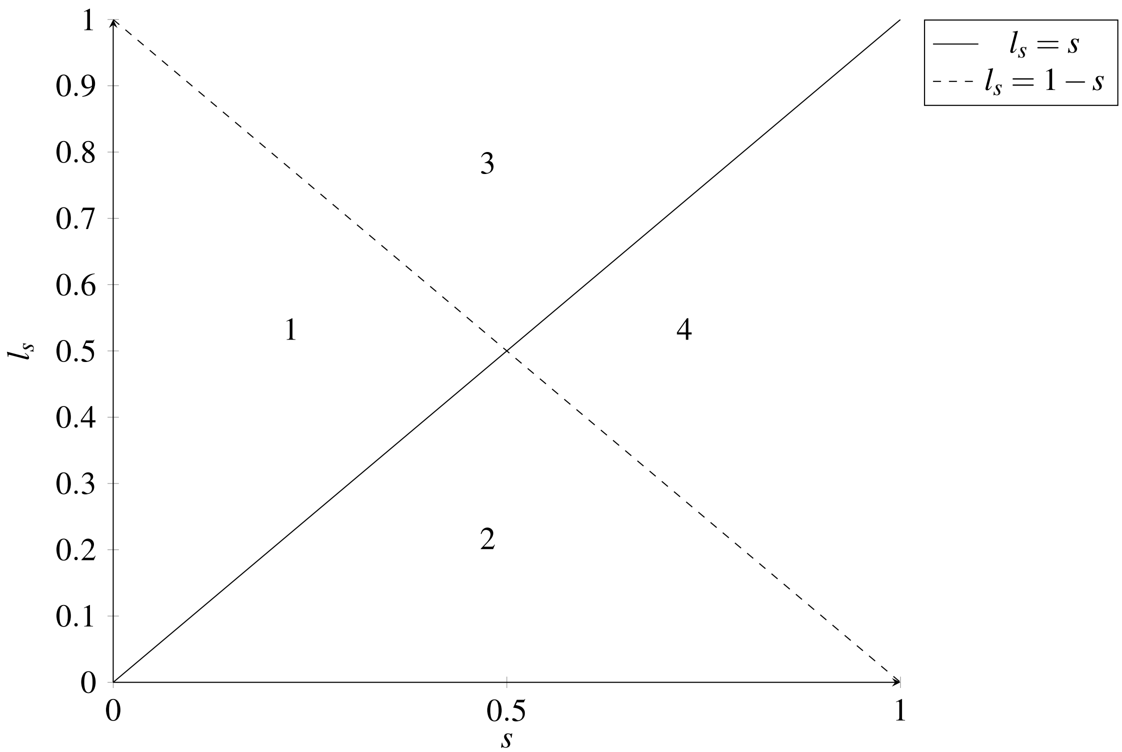

Figure 2: Four regions in the shock intensity and shock location space. Total displacement is different in each of these regions. Notice that the two lines cross at $s = 0.5$.

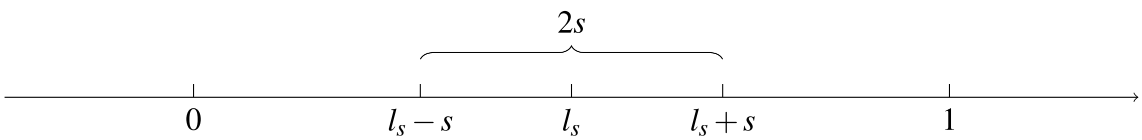

Figure 3: A shock of intensity s hits Country B at location l_s. Those in the interval $[l_s - s, l_s + s]$ are affected by the shock. Total displacement is $2s$.

Figure 4: A shock of intensity s hits Country B at location l_s. Those in the interval $[0, l_s + s]$ are affected by the shock. Total displacement is $l_s + s$.

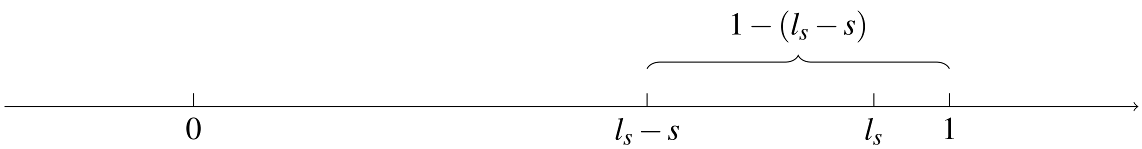

Figure 5: A shock of intensity s hits Country B at location l_s. Those in the interval $[l_s - s, 1]$ are affected by the shock. Total displacement is $1 - (l_s - s)$.

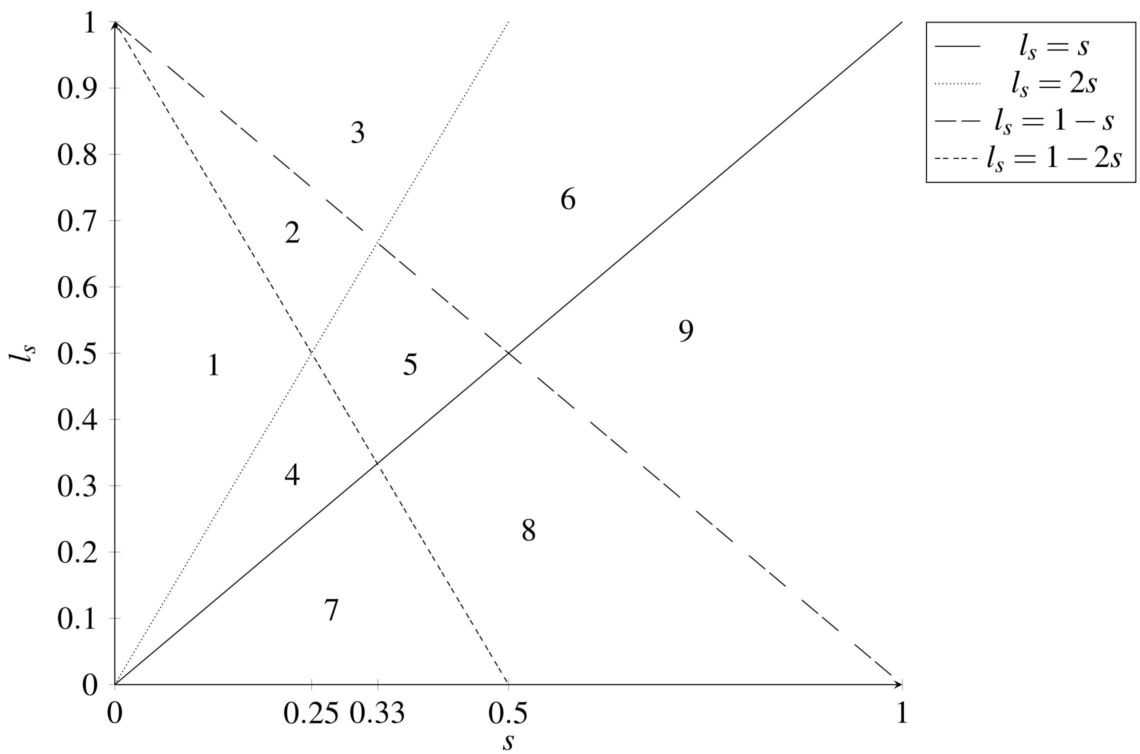

Figure 6: Areas of internal and external displacement

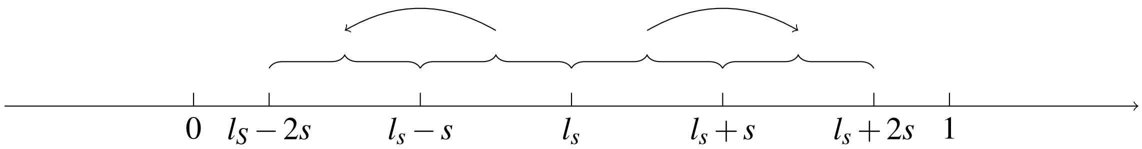

Figure 7: A shock of intensity s hits Country B at location l_s. Those in the interval $[l_s - s, l_s)$ move to $[l_s - 2s, l_s - s)$ and those in the interval $[l_s, l_s + s]$ move to $[l_s + s, l_s + 2s]$. There are no flows of refugees between Country B and either Country A or C. All displaced people remain as IDPs.

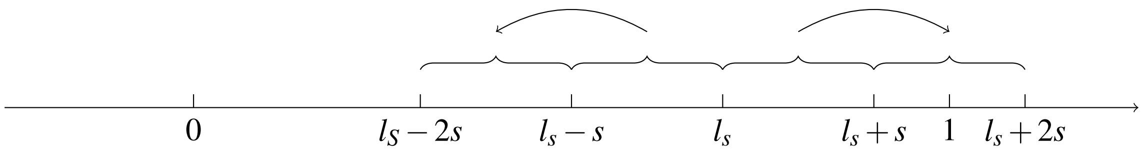

Figure 8: A shock of intensity s hits Country B at location l_s. Those in the interval $[l_s - s, l_s)$ move to $[l_s - 2s, l_s - s)$ remaining as IDPs. Those in the interval $[l_s, l_s + s]$ move to $[l_s + s, l_s + 2s]$. A fraction $l_s + 2s - 1$ of the population crosses the border to Country C, thus becoming refugees, and a fraction $1 - l_s - s$ of the population remains as IDPs.

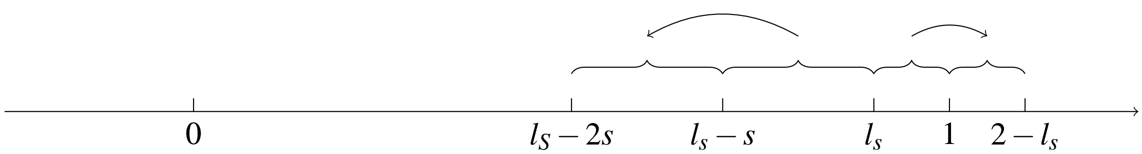

Figure 9: A shock of intensity s hits Country B at location l_s. Those in the interval $[l_s - s, l_s)$ move to $[l_s - 2s, l_s - s)$ remaining as IDPs. Those in the interval $[l_s, 1]$ move to $[1, 2 - l_s]$. A fraction $1 - l_s$ of the population crosses the border to Country C, thus becoming refugees.

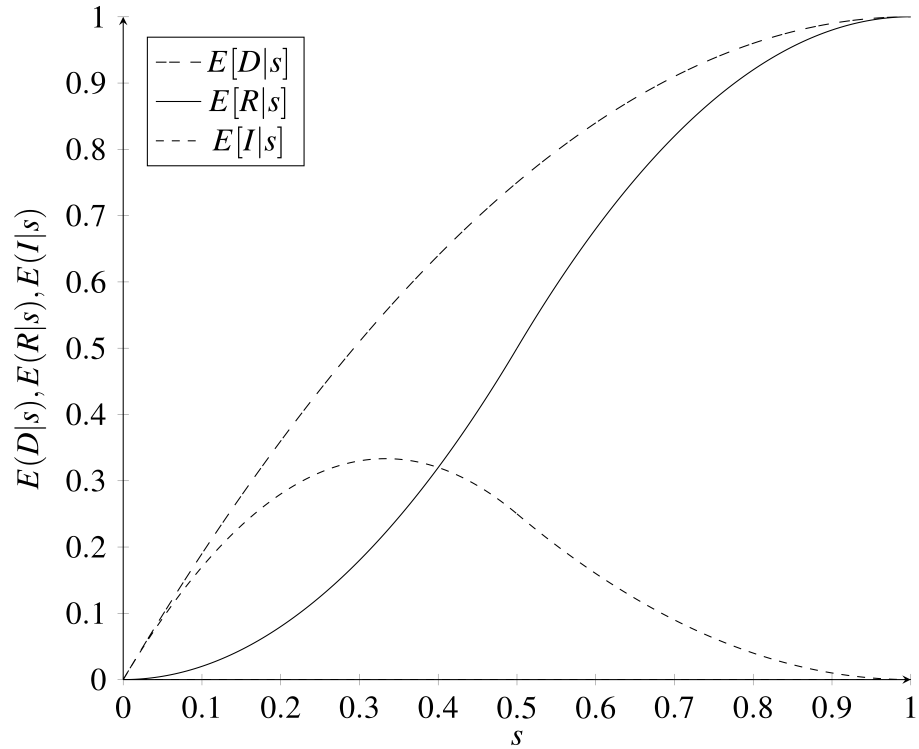

Figure 10: Expected values of groups of people

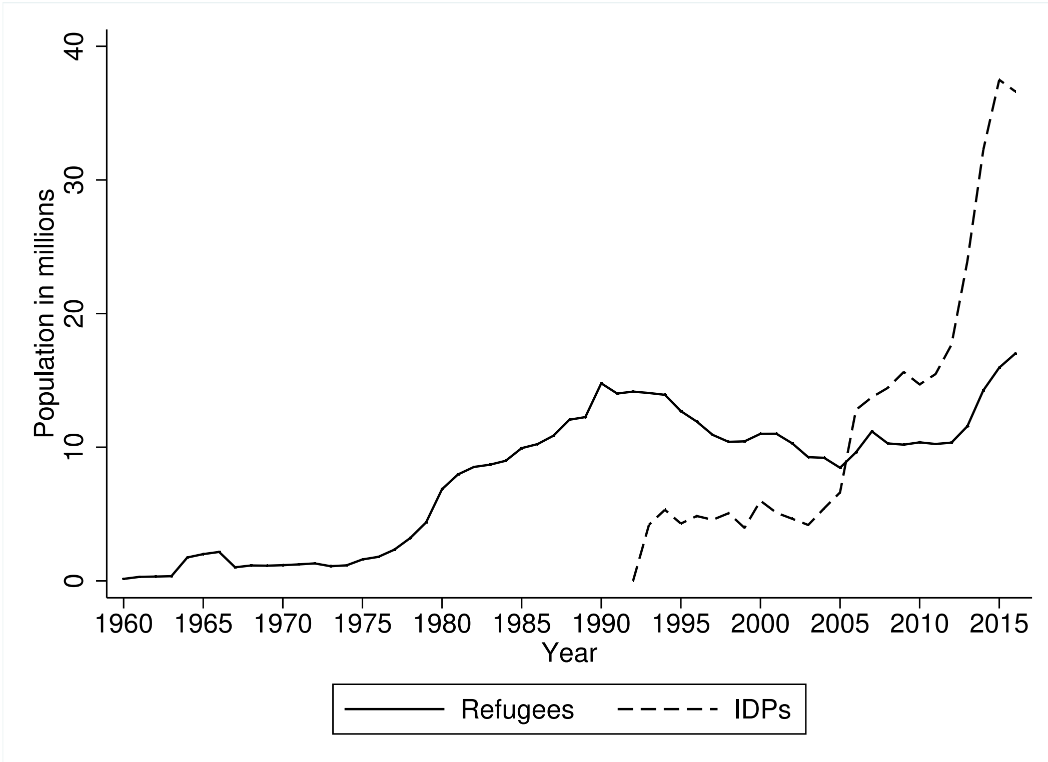

Figure 11: Trends in the aggregate number of refugees and IDPs

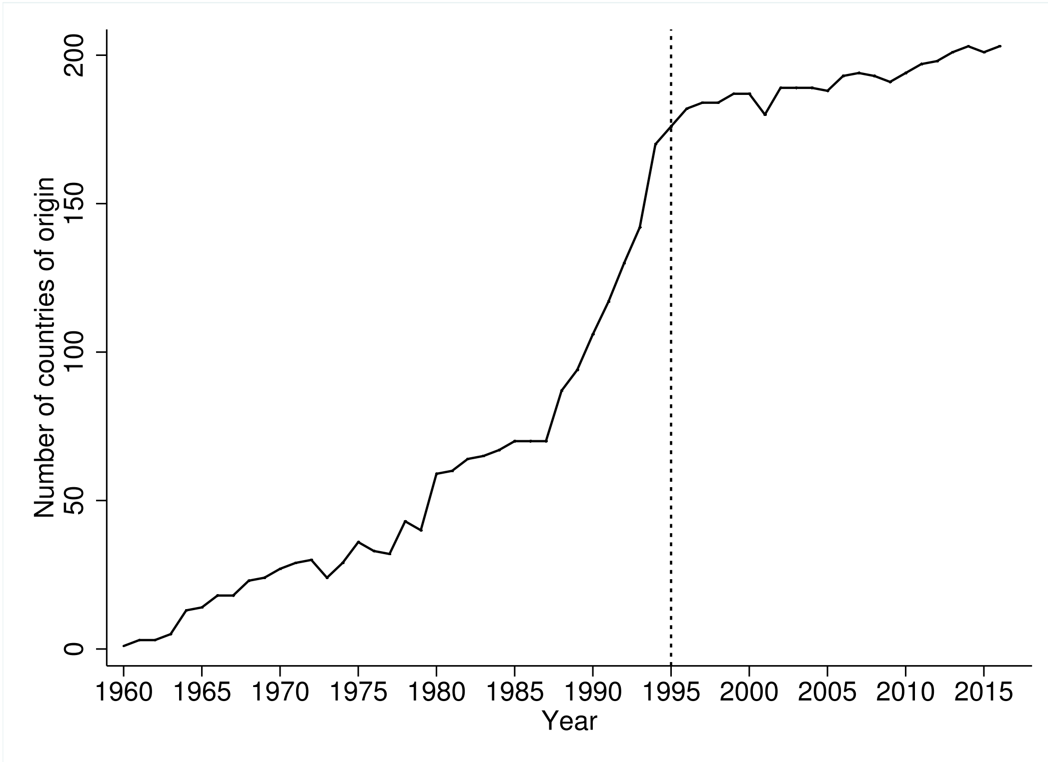

Figure 12: Number of countries listed as origin of refugees in the UNHCR population statistics. The vertical line divides the pre-sample (1960-1994) and sample periods (1995-2016).

Table 1: List of Countries

Countries and territories listed in the UNHCR population statistics included in the analysis:
Afghanistan, Albania, Algeria, Andorra, Angola, Antigua and Barbuda, Argentina, Armenia, Aruba, Australia, Austria, Azerbaijan, Bahamas, Bahrain, Bangladesh, Barbados, Belarus, Belgium, Belize, Benin, Bermuda, Bhutan, Bolivia, Bosnia and Herzegovina, Botswana, Brazil, British Virgin Islands, Brunei, Bulgaria, Burkina Faso, Burundi, Cabo Verde, Cambodia, Cameroon, Canada, Cayman Islands, Central African Republic, Chad, Chile, China, Colombia, Comoros, Congo, Costa Rica, Croatia, Cuba, Cyprus, Czech, North Korea, Denmark, Djibouti, Dominica, Dominican Republic, Ecuador, Egypt, El Salvador, Equatorial Guinea, Eritrea, Estonia, Ethiopia, Fiji, Finland, France, French Polynesia, Gabon, Gambia, Georgia, Germany, Ghana, Gibraltar, Greece, Grenada, Guatemala, Guinea, Guinea-Bissau, Guyana, Haiti, Honduras, Hong Kong, Hungary, Iceland, India, Indonesia, Iran, Iraq, Ireland, Israel, Italy, Ivory Coast, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Kiribati, South Korea, Kuwait, Kyrgyzstan, Laos, Latvia, Lebanon, Lesotho, Liberia, Libya, Lithuania, Luxembourg, Macao, Macedonia, Madagascar, Malawi, Malaysia, Maldives, Mali, Malta, Marshall Islands, Mauritania, Mauritius, Mexico, Micronesia, Moldova, Mongolia, Morocco, Mozambique, Myanmar, Namibia, Nauru, Nepal, Netherlands, New Caledonia, New Zealand, Nicaragua, Niger, Nigeria, Norway, Oman, Pakistan, Palau, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Portugal, Puerto Rico, Qatar, Romania, Russia (Soviet Union), Rwanda, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, Samoa, San Marino, Sao Tome and Principe, Saudi Arabia, Senegal, Serbia, Seychelles, Sierra Leone, Singapore, Slovakia, Slovenia, Solomon Islands, Somalia, South Africa, Spain, Sri Lanka, Sudan, Suriname, Swaziland, Sweden, Switzerland, Syria, Tajikistan, Tanzania, Thailand, Timor-Leste, Togo, Tonga, Trinidad and Tobago, Tunisia, Turkey, Turkmenistan, Turks and Caicos Islands, Tuvalu, Uganda, Ukraine, United Arab Emirates, United Kingdom, United States of America, Uruguay, Uzbekistan, Vanuatu, Venezuela, Vietnam, Yemen, Zaire, Zambia, Zimbabwe.
Countries and territories listed in the UNHCR population statistics not included in the analysis:
American Samoa, Anguilla, Bonaire, Cook Islands, Curaçao, French Guiana, Guadeloupe, Guam, Liechtenstein, Martinique, Monaco, Montenegro, Montserrat, Niue, Norfolk Island, Palestine, Saint Maarten (Dutch part), Saint Pierre et Miquelon, Svalbard and Jan Mayen, South Sudan, The Holy See, Tibetan, Wallis and Futuna Islands, Western Sahara.

Table 2: Exponential regressions: refugees

	(1)	(2)	(3)	(4)
(Log) Area	-0.4904***	-0.3153***	-0.3405***	-0.3408***
	(0.0814)	(0.0943)	(0.0688)	(0.0662)
(Log) Ruggedness	0.3422	-0.2569*	-0.3407**	-0.3784**
	(0.2526)	(0.1428)	(0.1544)	(0.1883)
(Log) Proximity	0.1169***	0.1017***	0.0823**	0.0875**
	(0.0419)	(0.0393)	(0.0400)	(0.0421)
Low intensity conflict	0.5232	0.4658*	0.4262**	0.4034*
	(0.3715)	(0.2708)	(0.2145)	(0.2174)
High intensity conflict	2.7842***	1.4917***	0.8885***	0.8442***
	(0.4277)	(0.3826)	(0.2086)	(0.2133)
Civil Liberties level 2			2.0607***	2.0685***
			(0.4322)	(0.4324)
Civil Liberties level 3			2.0913***	2.1427***
			(0.4543)	(0.4756)
Civil Liberties level 4			2.7370***	2.7539***
			(0.5185)	(0.5443)
Civil Liberties level 5			3.0028***	3.0490***
			(0.5868)	(0.6172)
Civil Liberties level 6			3.3458***	3.4084***
			(0.5667)	(0.6008)
Civil Liberties level 7			3.9158***	3.9849***
			(0.6600)	(0.7246)
(Log) GDP per capita			-0.3777**	-0.3915**
			(0.1777)	(0.1855)
(Log) # common ethnicity			0.1095	0.0983
			(0.1125)	(0.1093)
(Log) # colonial relations			0.1018	0.1094
			(0.3295)	(0.3329)
% desert > 20 %				-0.5911
				(0.6506)
Landlocked				0.0382
				(0.3344)
Pre-Sample Mean		0.4631***	0.4851***	0.4786***
		(0.0908)	(0.0485)	(0.0497)
Constant	-7.8443***	-3.9968***	-2.6189	-2.6316
	(0.8706)	(1.0009)	(2.4906)	(2.6564)
Pseudo - R^2	0.4695	0.7596	0.8258	0.8269
Number of observations	4,395	3,691	3,455	3,455
Number of countries	200	168	161	161

Lagged population used as exposure in all regressions which include year and continent dummies. Country-clustered robust standard errors in parentheses. One, two and three stars stand for 10, 5 and 1 per cent significance levels.

Table 3: Exponential regressions: IDPs

	(1)	(2)	(3)	(4)
(Log) Area	-0.6700***	-0.5265***	-0.4660***	-0.4864***
	(0.1167)	(0.1320)	(0.1319)	(0.1264)
(Log) Ruggedness	-0.1353	-0.5013***	-0.6382***	-0.6126***
	(0.1772)	(0.1877)	(0.2069)	(0.2189)
(Log) Proximity	0.2325***	0.2056***	0.2259***	0.2317***
	(0.0625)	(0.0581)	(0.0784)	(0.0775)
Low intensity conflict	2.2796***	2.1725***	1.6122***	1.6423***
	(0.5084)	(0.5363)	(0.4776)	(0.4684)
High intensity conflict	3.7109***	3.1088***	2.0410***	2.0134***
	(0.5821)	(0.6249)	(0.6345)	(0.6306)
Civil Liberties level 2			0.1799	-0.0497
			(1.5346)	(1.5198)
Civil Liberties level 3			1.9730*	1.6507
			(1.1813)	(1.1878)
Civil Liberties level 4			2.5737**	2.2679*
			(1.2014)	(1.2030)
Civil Liberties level 5			2.4682**	2.2481*
			(1.2539)	(1.2745)
Civil Liberties level 6			2.5747*	2.2566*
			(1.3319)	(1.3183)
Civil Liberties level 7			3.4573***	3.0998**
			(1.3349)	(1.3512)
(Log) GDP per capita			-0.5125*	-0.7030**
			(0.3007)	(0.3119)
(Log) # common ethnicity			0.1874	0.1886
			(0.1719)	(0.1841)
(Log) # colonial relations			0.1733	0.2203
			(0.2859)	(0.2834)
% desert > 20 %				-0.5571
				(0.8136)
Landlocked				-0.8236
				(0.6076)
Pre-Sample Mean		0.2708***	0.2992***	0.3232***
		(0.0604)	(0.0603)	(0.0573)
Constant	-10.8868***	-8.3968***	-4.0450	-1.9149
	(1.1515)	(1.2356)	(4.0938)	(4.0484)
Pseudo - R^2	0.5702	0.6704	0.6988	0.7108
Number of observations	4,395	3,691	3,455	3,455
Number of countries	200	168	161	161

Lagged population used as exposure in all regressions which include year and continent dummies. Country-clustered robust standard errors in parentheses. One, two and three stars stand for 10, 5 and 1 per cent significance levels.

Table 4: Exponential regressions: total displacement

	(1)	(2)	(3)	(4)
(Log) Area	-0.5852***	-0.4046***	-0.3970***	-0.4105***
	(0.0975)	(0.1059)	(0.0852)	(0.0801)
(Log) Ruggedness	0.1133	-0.3564**	-0.4557***	-0.4377**
	(0.2048)	(0.1489)	(0.1674)	(0.1812)
(Log) Proximity	0.1636***	0.1403***	0.1339***	0.1434***
	(0.0507)	(0.0403)	(0.0430)	(0.0425)
Low intensity conflict	1.4362***	1.3853***	1.0883***	1.0843***
	(0.4243)	(0.4216)	(0.3233)	(0.3141)
High intensity conflict	3.1726***	2.3605***	1.4707***	1.4320***
	(0.3818)	(0.4185)	(0.3843)	(0.3683)
Civil Liberties level 2			0.9272	0.8299
			(1.0391)	(1.0420)
Civil Liberties level 3			2.0201**	1.8588*
			(0.9441)	(0.9530)
Civil Liberties level 4			2.7725***	2.6221***
			(0.9173)	(0.9242)
Civil Liberties level 5			2.8663***	2.7531***
			(0.9622)	(0.9713)
Civil Liberties level 6			3.0599***	2.9244***
			(0.9839)	(0.9812)
Civil Liberties level 7			3.9046***	3.7649***
			(1.0253)	(1.0367)
(Log) GDP per capita			-0.4234**	-0.5496***
			(0.1991)	(0.2033)
(Log) # common ethnicity			0.1803	0.1816
			(0.1274)	(0.1290)
(Log) # colonial relations			0.0263	0.0714
			(0.2165)	(0.2150)
% desert > 20 %				-0.5094
				(0.5900)
Landlocked				-0.4412
				(0.3352)
Pre-Sample Mean		0.3380***	0.3555***	0.3671***
		(0.0616)	(0.0449)	(0.0479)
Constant	-8.4598***	-5.5005***	-3.2830	-1.8899
	(0.9494)	(0.9710)	(2.7763)	(2.7361)
Pseudo - R^2	0.5189	0.7067	0.7616	0.7676
Number of observations	4,395	3,691	3,455	3,455
Number of countries	200	168	161	161

Lagged population used as exposure in all regressions which include year and continent dummies. Country-clustered robust standard errors in parentheses. One, two and three stars stand for 10, 5 and 1 per cent significance levels.