Cities 95 (2019) 102407

Contents lists available at ScienceDirect

Cities

journal homepage: www.elsevier.com/locate/cities

A gravity model analysis of forced displacement in Colombia

Juan Francisco Saldarriaga*, Yuan Hua

Center for Spatial Research, Graduate School of Architecture, Planning and Preservation, Columbia University, 400 Avery Hall, 1172 Amsterdam Avenue, New York, NY 10027, United States of America

ARTICLE INFO

Keywords:
Forced displacement
Social network
Violence
Gravity model
Colombia
Latin America

ABSTRACT

Between 1985 and 2016 more than 7 million people were victims of forced displacement in Colombia. At the height of the conflict, more than 90% of municipalities in the country saw some form of displacement. In this study we extend the traditional gravity model of migration to analyze the flows of internally displaced people in Colombia between 1986 and 2015, and identify some of the main factors involved in people's choice of destination. We find violence at the origin municipality to be the most important driver of displacement. Similarly, destination municipalities with larger social networks of victims of displacement, larger overall populations, and closer to places of origin attract more displaced people. We propose forced displacement be treated as regional-level phenomenon and planners, city officials, and aid organizations focus their attention on medium-sized regional centers. Because of the importance of social networks in driving people's choice of destination, city officials, planners and aid organizations should closely collaborate with grassroots community organizations to adequately allocate resources and plan for new arrivals.

1. Introduction

Over the past five decades, more than 7 million people have been victims of death threats, massacres, kidnapping, torture, property loss, forced displacement or forced recruitment in the Colombian conflict. Amongst these crimes, forced displacement makes up the largest category: according to government records (Red Nacional de Información, 2016), between 1985 and 2016, approximately 6.9 million people were forced to leave their homes. This accounts for almost 88% of all the recorded victims of the Colombian conflict.

During this time period the vast majority of forced displacement occurred to or from urban agglomerations: 65% of all displaced people were forced out of municipalities with more than 20,000 inhabitants, and more than 80% of displaced people sought refuge in similar sized cities. More than half of all victims of displacement went to cities of more than 50,000 people and almost a quarter of them went to cities with more than 500,000 inhabitants.

As the number of people who suffered forced displacement increased exponentially, some cities saw their population swell dramatically, while others were depleted, often more than once. In 2002, for example, Tibú, a small municipality of 35,000 people in the northeast of Colombia, suffered almost 9000 forced displacements, which represents almost 25% of its population. Conversely, in that same year, the city of Quibdó, in the west of the country, received more than 25,000 victims of forced displacement, effectively increasing its size by more than 20% in a single year.

Maps, such as the one in Fig. 1, suggest displaced people travel mainly to large cities such as Bogotá, Medellín or Cali, or to towns near their place of origin. However, not all displacement follows the same pattern. In this study, we use *migration behavior modeling* and *regression analysis* to better understand what are the major elements associated with forced displacement in Colombia, and identify significant factors correlated with the choice of destination by displaced people. Specifically, we use *gravity models* to examine the correlation between a victims' path and the characteristics of the origin and destination municipalities.

Understanding the factors associated with the choice of destination for victims of forced displacement is crucial, not only in the context of the Colombian conflict, but for cities, planners, officials, and aid organizations all around the world. According to the Office of the United Nations High Commissioner for Refugees (UNHCR), in 2018 alone there were 2.3 million new displaced people, bringing the worldwide total to more than 70 million. This includes 25.9 million refugees, as well as 41.3 million internally displaced people and 3.5 million asylum-seekers (Office of the United Nations High Commission for Refugees, 2019).

Many of these refugees and internally displaced people live in purposely built camps but a large number of them reside in cities. For example, the Turkish Ministry of Interior estimates that more than

* Corresponding author.
E-mail addresses: jfs2118@columbia.edu (J.F. Saldarriaga), yh2876@columbia.edu (Y. Hua).

https://doi.org/10.1016/j.cities.2019.102407
Received 31 July 2018; Received in revised form 26 June 2019; Accepted 12 July 2019

Fig. 1. Internally displaced people in Colombia (1900–2016).

500,000 Syrian refugees live in Istanbul, and more than 400,000 live in Gaziantep and Hatay respectively. For the latter two, this represents more than 20% of their current population (Republic of Turkey, Ministry of Interior, 2019).

In addition, as climate change intensifies cities around the globe will see a dramatic increase in the number of displaced people seeking refuge. In Vietnam alone, the Internal Displacement Monitoring Center estimates that in 2017 ten natural disasters caused more than 600,000 displacements, many towards Ho Chi Minh city (Internal Displacement Monitoring Centre, 2019). According to the Internal Displacement Monitoring Centre “disasters, climate change impacts and conflict trigger displacement both to and within urban areas. A study conducted in Bangladesh suggests that a significant proportion of people who live in informal urban settlements may have been displaced from rural areas by riverbank erosion, a major hazard in the country, projected to increase in the coming years […] The Bangladeshi capital, Dhaka, has been identified as the country's main destination for people fleeing disasters and climate change impacts, and local authorities have been unable to cope with the influx.” (2019) Globally, the number of disaster related displacements in 2018 is estimated to have been more than 17 million (Internal Displacement Monitoring Centre, 2019).

In terms of housing, health care, equal rights, education, and economic opportunities, refugees and internally displaced people are as disadvantaged as the urban poor, and require assistance (Crisp, Morris, & Refstie, 2012). In addition, urban crisis might trigger further rounds of forced displacement, affecting these already highly vulnerable populations. It is therefore imperative for cities, planners, government officials, and aid organizations to better understand the factors that drive people's choice of destination when forced to relocate. A better understanding of this phenomenon will help cities formulate more adequate policies and plans, secure resources, and provide services.

Previous studies around forced displacement in Colombia have mostly focused on the causes, consequences, and characteristics of this phenomenon. Kalyvas (2006), Engel and Ibáñez (2007), and Steele (2009) highlight forced displacement as a war strategy used to strengthen military control, prevent defection and expropriate properties and resources. Ibáñez (2009) further concludes that forced displacement in Colombia has been provoked by “attacks by armed groups, indiscriminate violence, or the mere presence of armed groups”. She finds that, contrary to other countries, in Colombia “people migrate individually, and few move outside the country's borders”.

Ibáñez and Moya (2006), Ibáñez and Vélez (2008) as well as Ibáñez and Velásquez (2009) identify the following as consequences of the massive displacement that has occurred in the country: destruction of social networks, disintegration of families, high unemployment rates, poor socio-economic conditions in many of the destination cities, as well as loss of economic, financial, physical, human, and social capital throughout the country. Using survey data from 2322 displaced households, Ibáñez and Moya (2006) found that “most displacement (86.2 percent) is a reaction to being a victim of direct attacks”, that the majority of displaced people (76.1%) moved individually, and that more than 88% moved directly to the destination municipality.

However, despite the large amount of research around forced displacement, few studies have looked at what factors determine a victim's choice of destination, specially from an urban perspective. Ibáñez and Moya (2006) found that “potential support from family, friends, and government programs, as well as distance from origin and anonymity were the main reasons for choosing destination sides”. More specifically, “support from family and friends is identified as the main reason for the selection of a destination site (65.3 percent of households), while potential access to government programs for the displaced population seems to play a less crucial role (26.3 percent)” (Ibáñez & Moya, 2006).

Steele (2009) highlighted that victims who were collectively victimized, generally have three options when choosing destinations: moving to a rival group's stronghold, clustering with other victims of similar crimes or origins, or seeking anonymity in larger communities. Through analysis of conflict-derived interregional migration based on the gravity models, Lozano-Gracia, Piras, and Ibáñez (2010) noted that “violence appears to be one of the most relevant pushing effects together with the absence of institutions and the dissatisfaction with the provision of basic needs”, while on the destination side, “most populated regions are more attractive as well as areas with a sufficient level of fulfillment of basic needs”.

In other contexts, such as interregional educational or economic migration, studies point to place amenities, work opportunities, and lifestyle as significant factors in people's choice of destination (Darchen & Tremblay, 2010; Frenkel, Bendit, & Kaplan, 2013 and Lawton, Murphy, & Redmond, 2013). Kaplan et al. found that “social networks, norms and lifestyle are main factors in forming inter-regional migration.” (2016) And in the case of technical workers in Denmark, Dahl and Sorenson (2010) found the following factors in order of importance to be considered when choosing where to work: proximity to the worker's current residence, proximity to their parents, the number of high school and college classmates in the region, proximity to places they have lived in previously, and income. “For the typical Danish scientist, engineer or medical worker, social factors swap economic considerations.” (Dahl & Sorenson, 2010).

Following Lozano-Gracia et al.'s research approach, this project uses gravity models and regression analysis to study forced displacement in Colombia. Gravity models are one of the most prominent mathematical models used to understand social phenomena and their spatial interactions, including migration, commuting, trade, and information exchange. They were first proposed in E.G. Ravenstein's “The Laws of Migration” (1885, 1889), which borrowed Newton's laws of gravity to explain the gravity-like properties of internal migration flows (Alhanaee & Csala, 2015; Greenwood, 2005; Poor, Alimi, Cameron, & Maré, 2016).

In their original and most commonly applied form, migration from region ‘i’ to region ‘j’ is positively correlated to the population size at origin and at destination, and is inversely correlated to distance between the two. However, gravity models have continually been modified and extended to include other variables representing characteristics of host or source locations such as language, economic conditions, unemployment rates, or political instability (e.g. Greenwood, 1975; Borjas, 1987, 1989; Schmeidl, 1997, Karemera, Oguledo, & Davis, 2000). Some notable cases in which gravity models have been used to analyze migration include intercity movement of people in the U.S. (Zipf, 1946), international migration to North America (Karemera et al., 2000), internal and international migration in New Zealand (Poor et al., 2016) and Salvadoran migration to the United States (Stanley, 1987).

As in some of the previous cases, in analyzing forced displacement other explanatory variables should be included. In addition to population and distance, Lozano-Gracia et al. (2010) introduce violence, number of migrants previously arrived at the destination, income levels, service institutions and neighbors' characteristics into their model. In a different study, Alhanaee and Csala (2015) analyzed Syrian refugee movement through the number of institutions providing education and protection.

Our study coincides with Lozano-Gracia et al.'s analysis of Colombian forced displacement in terms of its purpose and main methods: both attempt to identify the determinants of movement and used gravity models, taking into consideration the intensity of crimes and regional effects. However, our study differs from theirs and improves upon the current body of research in four main ways: first, the breath and granularity of our data is exponentially higher allowing us to perform our analysis at a micro-scale (municipalities as opposed to regions or countries); second, we include different variables and employ multiple ways of calculating the distance between origin and destination, which allow us to perform a more nuanced analysis; third, instead of using the aggregate number of displaced people between 2001 and 2006, we expand the range to 1986 to 2015 and conduct regression analysis based on both annual and aggregate data, performing a more fine-grained analysis of the dynamics and evolution of forced migration for those 30 years of conflict; and fourth, we conduct regressions using subsets of the data in order to investigate whether different types of displacement or different demographic characteristics of the victims, have different effects on the choice of destination.

Our results reveal that intensity of violence at the origin municipality is a significant factor associated with outbound displacement. Similarly, distance between origin and destination, and population size at the destination, are statistically correlated to the number of forcibly displaced people arriving at a municipality. These findings are in line with previous studies. However, *we found the number of previously displaced people from the same origin at the destination municipality to be the strongest predictor of displacement volume.* This confirms qualitative survey data from other studies (Ibáñez and Moya, 2006) and suggests a much stronger link between social networks of victims and their choice of destination. While still in line with some of the other studies of economic or educational migration (Dahl & Sorenson, 2010; Kaplan, Grünwald, & Hirte, 2016), the strength of this relationship has not been previously documented in relation to forced displacement.

2. Methodology

2.1. Model

This project borrows the gravity model from classic migration models to estimate the influence of the source and destination on the size and orientation of forced displacement flows in Colombia between 1986 and 2015. Besides the classical gravity model variables, such as population and distance, we take into consideration the intensity of violence at the origin and destination municipalities as well as in the origin and destination regions. Similarly, we consider the level of community participation at the origin and destination. Furthermore, since displaced people tend to cluster in space (Bartel, 1989; Lozano-Gracia et al., 2010), we introduced a variable representing extent of the social network at the destination.

Given these variables, the reformulated equation for our gravity model is as follows:

$$M = f(P, D, V_{site}, V_{region}, C, N),$$

where displaced population (M) is a function of population (P), distance (D), intensity of violence at the municipality and region (V_{site} and V_{region}), community participation (C) and social network of migrants (N).

2.2. Data

The main dataset used in this project comes from the Registro Único de Víctimas, the official database of victims of the conflict in Colombia. It was assembled by the Unidad para la Atención y Reparación Integral a las Víctimas - the Colombian government agency in charge of aiding the victims of the conflict and providing reparations - and is based on self-reported accounts by the victims of the conflict.

Specifically, the dataset contains reports on victims of the following crimes: terrorism, threats, sexual violence, forced disappearance, forced displacement, homicide, explosion of landmines, bombs or munition, kidnapping, torture, recruitment of minors, forced abandonment of land, and loss of property. Each victim also provided their gender, ethnicity and date of birth. Table 1 displays the number of reports per type of crime. Victims also specified the armed group that committed the crime, including guerrillas, paramilitaries, Fuerza Pública (the Colombian armed forces), Bacrim (new criminal bands, often an offshoot of the paramilitaries), other or no-identified, and victims of forced displacement stated the municipality they were displaced from and the municipality they migrated to.

The advantage of this dataset lies in its magnitude, granularity, and time span. As of February 2016, the dataset contained approximately 9.6 million records of crimes, 7.9 million individual victims, and covered most of the 1122 municipalities in Colombia, tracing back mostly from 1985 to 2016. Of the crimes recorded, approximately 8.7 million (90%) are of direct victimization and of these, 7.7 million (88.8%) of direct forced displacement. The dataset is not yet complete and the Colombian government is still actively collecting information on crimes and victims.

Nevertheless, this dataset has some limitations. First, there is little evidence that it includes all victims of the Colombian conflict. According to Ibáñez and Velásquez (2009), "a little more than 70% of displaced households registered in RUPD [Registro Único de Población Desplazada], with close to 8% not registering voluntarily and roughly 15% being overtly excluded". Most of the non-registered victims lack knowledge of the program. In addition, distrust in government, lack of accessibility to attention centers, or fierce combats can prevent a victim from registering in the database. In addition, although government officials claim to have validated every record, some fraudulent claimants pretending to be victims in order to receive program benefits have been identified in the dataset. (Ibáñez & Velásquez, 2009). This being said, recent analysis released in June 2017 by the attorney general (Fiscalia General de la Nación) has the highest estimate of false reports at no more than 22,000, which represents less than 0.3% of the whole dataset (Gossaín, 2017).

Finally, in working with this dataset we found some erroneous records. For example, between 1986 and 2015, some victims' date of birth was later than the date of the crime and some forced displacement victims arrived at their destination before the crime occurred. However, these errors only account for approximately 2.8% of the dataset and have been excluded from our analysis.

Table 1
Distribution of types of crimes: direct and indirect (1900–2016).

Crime	Direct victims (%)	Indirect victims (%)	Total (% of total)
Terrorism/combats/skirmishes	75,635 (0.90%)	1363 (0.16%)	76,998 (0.83%)
Threats	343,065 (4.07%)	0 (0%)	343,065 (3.70%)
Sexual violence	10,978 (0.13%)	2426 (0.29%)	13,404 (0.14%)
Forced disappearance	37,992 (0.45%)	107,944 (12.94%)	145,936 (1.57%)
Forced displacement	7,540,682 (89.41%)	0 (0%)	7,540,682 (81.36%)
Homicide	253,012 (3.00%)	714,166 (85.58%)	967,178 (10.44%)
Explosion of landmines/bombs/munition	10,538 (0.12%)	2154 (0.26%)	12,692 (0.14%)
Kidnapping	26,802 (0.32%)	3858 (0.46%)	30,660 (0.33%)
Torture	4721 (0.06%)	1931 (0.23%)	6652 (0.07%)
Recruitment of minors	5853 (0.07%)	484 (0.06%)	6337 (0.07%)
Forced abandonment of land	10,882 (0.13%)	0 (0%)	10,882 (0.12%)
Loss of property	113,862 (1.35%)	168 (0.02%)	114,030 (1.23%)
No data	21 (0.00%)	4 (0.00%)	25 (0.00%)
Total	8,434,043 (91.00%)	834,498 (9.00%)	9,268,541 (100%)

Table 2
Description of variables.

Variable	Description
Displacement (M)	Logarithm of number of displaced people from origin to destination
Population (P)	Logarithm of total population at origin and destination
Distance (D)	- Straight line distance in meters - Distance by road travel in meters - Time by road travel in seconds - from the center of origin to center of destination municipality
Violence intensity (V)	Logarithm of number of victims per 10,000 inhabitants at origin and destination
Regional violence intensity (V_r)	Logarithm of number of victims per 10,000 inhabitants within a 100 km radius of origin and destination (excluding violence specifically at origin or destination)
Community participation (C)	Percentage of people participating in community organizations (2005)
Previous displacement from same origin (N_s)	Logarithm of total number of displaced people at the destination from the same origin in the past one, two, three, four, five or ten years
Previous displacement from different origin (N_d)	Logarithm of total number of displaced people at the destination from different origins in the past one, two, three, four, five or ten years

Note: All variables are measured at the municipality level.

Besides the main dataset, we also used data from the Colombian National Administrative Department of Statistics (DANE). This included population data for 1985 to 2015 which contains projections based on the 2005 national census. In addition, we used the results of a 2005 survey on community organization and participation by the DANE. The level of community participation was based on the number of people who responded affirmatively to the question "did someone in this household participate actively in an organization that benefits the community?"

2.3. Main variables

The main variables used in the different regressions are listed in Table 2. For a better fit and easier interpretation of the regression coefficients as elasticities, all variables except the percentage of community participation were transformed using the natural logarithm. To be able to work with zero values in some of our variables (for example, zero previous arrivals at a specific municipality) we added 10% of the smallest value to all rows in our data.

This way, the specific coefficients can be interpreted as follows: change of 1%age point in the independent variable results in change of x percent in the dependent variable, with x being the coefficient for the independent variable. More precisely, the log-log regression formulation is as follow:

$$\ln(y) = \beta_0 + \beta_1 \ln(x)$$

The interpretation of the coefficient is that if we change x by one percent, we'd expect y to change by β1 percent; namely, $\%\Delta y = \beta_1 \%\Delta x$.

Our dependent variable was the number of displaced people for each pair of origin and destination, for each year between 1986 and 2015, including the people migrating within the origin municipality. Over our 30-year study period, 32% of forced displacement occurred within the municipality of origin and 21% of victims never left their municipality, even after being victimized more than once. In these cases, the drawbacks of moving to a different municipality (loss of social network, distance to travel, property loss) probably outweighed the security benefits.

Our independent variables included population, distance, violence per capita, community participation and social capital at destination. Population is a common variable in gravity models: the size of municipality is widely considered to be positively correlated with the number of migrants - larger municipalities are expected to yield more outflow and migrants are expected to prefer more populated municipalities as destinations. In addition, we expected larger cities to be preferable for victims in search of anonymity (Lozano-Gracia et al., 2010; Steele, 2009).

Distance between origin and destination is also a common variable in gravity models. We expected the correlation between distance and number of displaced people to be negative, reflecting not only the travel and time costs but also the psychological cost moving far from your home town. In this study, we measured distance in three different ways: straight line distance ("as the crow flies"), shortest road travel distance, and time it would take to travel by road (assuming people drive). According to Poor et al. (2016) the distance measured as travel time is expected to be the best proxy of actual cost.

We measured the straight distance using the GIS software ArcMap, and distance by road and time by road using the Google Maps Directions API. Out of 99,080 pairs we excluded 9232 (approximately 9.3%) because Google Maps could not find a road connection between them (travel between some of these municipalities happens by water but calculating this travel distance and time falls beyond the scope of this project). Admittedly, the latter two distance measurements (road distance and time) are not necessarily 100% accurate, as the road network has surely changed since 1986. However, they do represent a good approximation to what people would have to travel and a good alternative to just using straight distance, especially in a country with such extreme topography as Colombia.

Our third independent variable, intensity of violence, can be defined as the number of crimes (in our dataset) for every 10,000 people. Since we focused on violence-caused forced displacement, we expected this variable to be one of the main determinants of the number of displaced people. We assumed that people tend to leave their municipalities as violence increases, and municipalities with lower violence will attract more displaced people. In this sense, intensity of violence at origin was expected to have a positive correlation to displacement, while the correlation at destination was expected to be negative.

We calculated two types of intensity of violence, one at the municipality level and one at the regional level (both for origin and destination). Both were expected to have similar correlations with the number of displaced people, but we believed the regional one to be weaker. Regions were defined using a 100 km buffer around each municipality. Regional violence was calculated without taking into account the violence or the population of the specific municipality, in order to not duplicate variables.

We assumed that the level of community participation in a specific municipality would indicate a more engaged and active population and thus discourage displacement. To test this hypothesis, we used an independent variable representing the percentage of people participating in community organizations (Departamento Administrativo Nacional de Estadísticas, 2005). Since this variable was already in percentage form, it was not necessary to convert it to logarithmic type. We assumed that a high level of community participation at origin will be negatively correlated to displacement and a high level of community participation at destination will be positively correlated with arrivals.

Table 3
Summary statistics of regression variables (1986–2015).

	Mean	Maximum	Minimum
Number of OD pairs	11,392	20,658 (2007)	2203 (1986)
Number of displacements	245,144	737,785 (2002)	12,549 (1986)

Finally, based on the assumption that victims of displacement would want to move where they already know people and cluster in space, we included an independent variable representing the social network for victims at destination. This variable was calculated by adding the number of displaced people from the same origin that arrived at the same destination in the previous years. We tested the same variable using previous arrivals in the past one, two, three, four, five and ten years and chose the one that yielded the best fit for our analysis as the final variable. Our assumption was that this social network provides support for the new arriving victims and in turn makes arriving at this location easier for victims of displacement. We expected to see a positive correlation between large arrivals in the past and arrivals in the present. In addition, we also included a variable representing arrivals in the past two years from different origins, in order to test if new arrivals are clustering specifically around displaced people from the same origin or victims of displacement in general, regardless of their origin.

2.4. Analysis

Our analysis was conducted using Pandas and NumPy (two data analysis libraries for the Python programming language) through the Jupyter Notebooks platform. To visualize the correlation between variables, we conducted correlation matrix analysis between variables through scatter plots and Seaborn Pairplots (SNS) Heatmaps using Seabon (a visualization library for Python).

To further measure the correlation, we ran regression analysis for each year between 1986 and 2015. We experimented with two models for the explanatory variables of population and intensity of violence: one using the data of the same year when the displacement happened, and another using the data of the year before the displacement happened. In other words, we wanted to see if the population or violence data of 1994 could explain the displacement that occurred in 1995, or if only the population and violence of that same year could explain it. We conducted both ordinary least squares (OLS) and generalized linear model (GLM) regressions. GLM regressions were also included because they allow dependent variables that have error distribution models other than normal distributions. The dependent variable (number of people displaced) is not linearly distributed: the larger the number, the lower the frequency. Hence, we used the Poisson distribution mode when conducting GLM regression. The regression models were run through StatsModels (a Python module for statistical analysis).

We ran regressions based on the basic gravity model (with explanatory variables of population and distance only) and on our extended equation. Considering the fact that some municipalities do not have other neighboring municipalities in the region, we conducted another regression for all pairs of origin and destination without regional variables. The three measures of distance and six measures of social network were also tested and compared. The measures that had the best fit for our model were used in our final regressions and further analysis.

We also ran regression based on the aggregate number of displaced people to avoid the potential biases that lie in annual observations. Displacement flows were aggregated over periods of five years (1986–1990, 1991–1995, 1996–2000, 2001–2005, 2006–2010, 2011–2015); population was averaged over every five years; intensity of violence at the municipality and region was summed over that period; and previous displacement was calculated as number of people arriving at the destinations n years before this period (the exact number years was determined by regressions run on annual data).

We conducted four more regressions using subsets of the dataset based on victims' characteristics and on the group that committed the crimes. The first regression centered around displacement towards other municipalities, as opposed to displacement within the same municipality. We wanted to understand if people who were determined to leave their municipality had different preferences in terms of destination. In the second subset, we reclassified the displacement flow by the reported gender of the victims. In the third, we divided the victims into three subsets based on their age at the moment of displacement: *"young"* for victims under the age of 18, *"adult"* for victims between 18 and 65 and *"elderly"* for victims over 65. Problematic records, like ones with missing date of birth, were removed from these regressions. Finally, displacement and intensity of violence at destination were subdivided based on the armed group that committed the crime, Guerrillas, Paramilitares, Fuerza Pública, Bacrim, other or non-identified. Through this reclassification we wanted to test if victims of displacement by a particular group would prefer destinations with low levels of violence caused by that same group.

Our regression took the following form:

$$\ln M_{OD} = \ln \beta_0 + \beta_1 \ln P_O + \beta_2 \ln P_D + \beta_3 \ln D_{OD} + \beta_4 \ln V_O + \beta_5 \ln V_D + \beta_6 \ln V_{rO} + \beta_7 \ln V_{rD} + \beta_{10} \ln C_O + \beta_{11} \ln C_D + \beta_8 \ln N_s + \beta_9 \ln N_d$$

The abbreviation for the variables follow the ones listed on Tables 1 and 2. The labels O and D indicate variables measured at origin and destination. The label rO refers to neighboring region at origin and rD to destination-centered region.

3. Results

Table 3 shows summary statistics for the sample and the dependent variable considered in the regression models. The number of observations refers to the sample size used in the regression, which is the number of origin-destination pairs that experienced forced displacement in any particular year. The number of displacements is the dependent variable: the total number of victims of forced displacement for each year.

Initial regressions show the explanatory variables for the same year as the displacement have a better fit with our gravity model than the ones from the previous year. In addition, the regressions based on travel time, as opposed to straight distance or road distance, and using the number of displacements in the previous two years had the highest R-square values. Hence, our analysis centers mainly on the results from the regressions with these three variables.

The following scatter plots (Fig. 2) illustrate the correlation between the number of displaced people for each origin-destination pair and the explanatory variables for 2005 (not in logarithmic form). From these plots we can deduce that distance (travel time) and rate of community participation are negatively correlated with the number of displaced people, while all other explanatory variables appear to be positively correlated.

The Correlation Matrix Heatmap (Fig. 3) displays the correlations between pairs of variables. The number of previous arrivers tends to have the strongest positive correlation to the number of people displaced, while distance from origin to destination tends to have the strongest negative one.

Table 4 reports the regression results based on annual data and aggregate data. The basic model refers to the regressions following the basic gravity model, involving only population and distance. The final model is based on our extended gravity model, including other variables. However, certain insignificant explanatory variables have been excluded from the final model. The R-squared for the OLS models suggest that close to 34% the variance in the number of displaced people of annual observations and 41% of aggregate data can be explained by population, distance, previous displacement, and intensity of

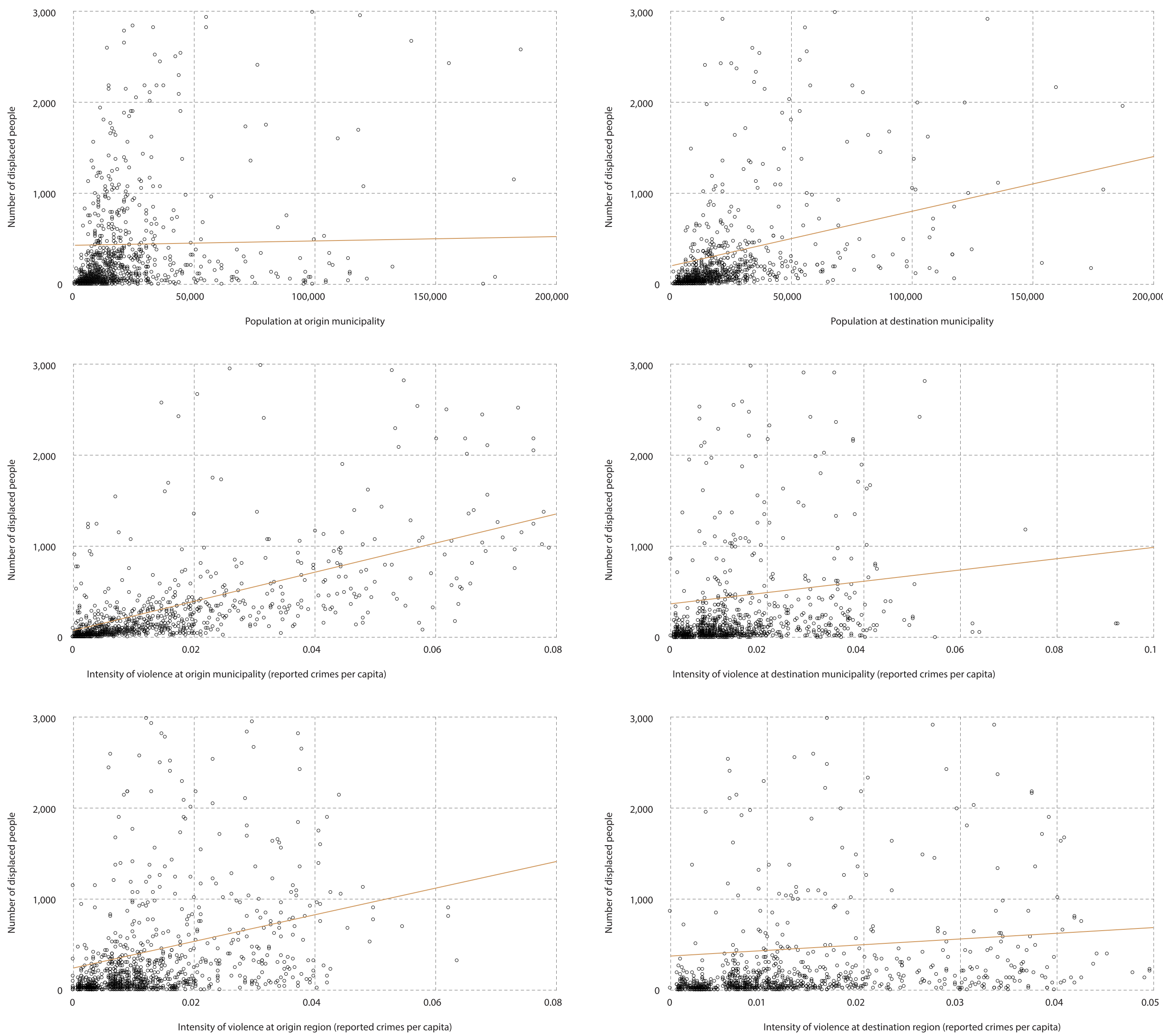

Fig. 2. Scatter plots with correlations between number of displaced people and explanatory variables (2005).

violence at origin municipality. The highest R-squared for an annual observation is 0.428 and for the aggregate data is 0.479.

There is a significant increase in the R-squared between the basic and the final models, which underlines the importance of the other variables in addition to distance and population. The direction of effect of the coefficients is relatively stable across specifications, indicating the reliability of annual observations and the regression models.

In all cases, the strongest predictor of displacement volume is the number of displaced people from the same origin in the previous two years. This variable's coefficients in both the OLS and GLM models are positive and statistically significant. This suggests that if the amount of displaced people that arrived to the destination municipality from a particular origin in the past two years is large, that municipality will keep on attracting high numbers of displaced people from that same origin. This confirms the idea that the social network and potential support from known people is particularly attractive to people suffering from forced displacement (Bartel, 1989; Lozano-Gracia et al., 2010).

Our results also point towards distance between origin and destination as a significant factor in a victim's choice of destination. As is shown on Table 4, the effect of distance, measured as travel time is negative and significant. Distance acts as a strong hindrance when people choose their destination. Larger displacement flows are directed to closer places.

Population size at the destination municipality and the number of people displaced from other origins in the previous two years are also significant factors in the flow of displaced people in Colombia. However, these correlations are not as strong as the two mentioned above. In accordance with the general migration literature, our results show that displaced people tend to choose more populated destinations. In addition, they also choose destinations that have received other displaced people, regardless of their place of origin. Nevertheless, this last preference is weaker than the one for municipalities with migrants from the same place of origin.

In addition, population size and intensity of violence at the origin municipality are also significant factors positively correlated with the number of displaced people.

Intensity of violence at the destination municipality, as well as at the neighboring origin and destination regions were also considered but

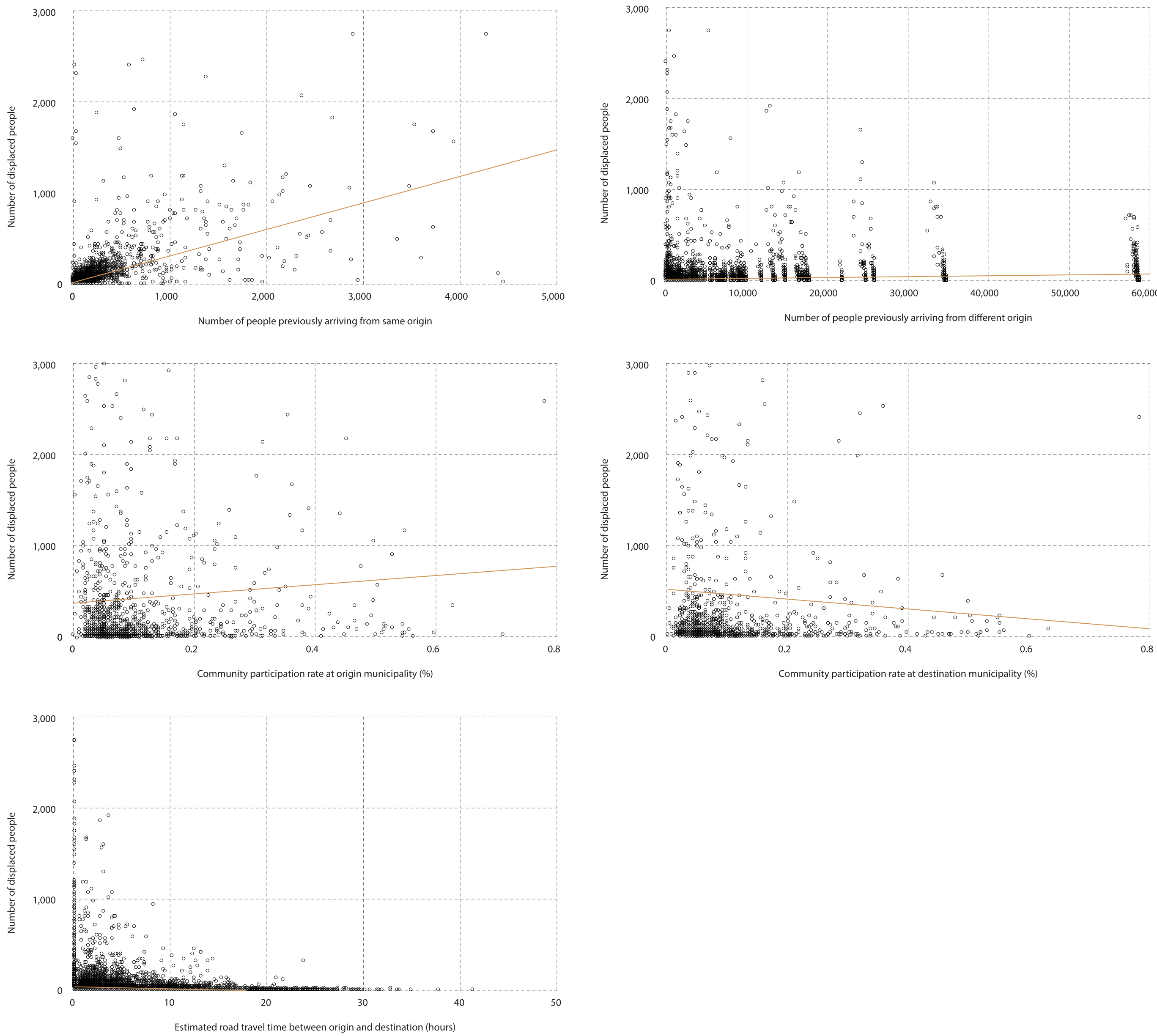

Fig. 2. (*continued*)

ultimately dropped from the final regression models. The coefficients for these variables were inconsistent and not always statistically significant.

Table 5 shows the results of the regression that include the community participation variable for 2005. The effect of the participation rate at the origin municipality is not significant in either the OLS or the GLM regressions. However, community participation rate at the destination is significant and positively correlated to the number of displaced people, indicating higher levels of community participation increase the pull effect of a destination.

Table 6 shows the distribution of the different subsets of the dataset: intra-municipality vs. inter-municipality displacement, gender of victim, age of victim, and armed group that committed the crime. Between 1986 and 2015, 31.5% of forced displacement occurred within the origin municipality; the percentage of male vs. female victims was evenly distributed; there were less elderly victims of forced displacement than young and adults; and the Guerrillas were the leading cause of displacement with 39.96% of the cases attributed to them. Since victims can report more than one actor as causing the displacement, the proportion of crimes attributed to the different actors of the conflict adds up to more than 100%.

The results of the regressions using the different subsets of the data are consistent with the main results using the complete dataset. Intra-municipality and inter-municipality displacement, as well as displaced people from different genders and age, and people displaced by different actors, appear to have similar considerations and preferences when choosing destination municipalities. All regressions point towards intensity of violence at origin, distance to destination, and presence of displaced people from the same origin at the destination, as the most important variables correlated with the choice of destination. In

Fig. 3. Correlation Matrix Heatmap (2005).

addition, victims of displacement did not seem to be affected in their choice of municipality if the violence at the destination is caused by the same armed group that displaced them.

4. Discussion

The results of our study match our previous expectations in multiple ways: more populated municipalities tend to generate more displacement, while larger population sizes at the destination municipalities attract more migrants; social networks, in the form of other displaced people from the same origin municipalities at the place of destination appear to be a crucial pull factor, drawing more migrants; displaced people tend to move to places closer to their origin municipalities and, everything else being equal, higher levels of violence at the origin tend to produce more displacement.

We expected higher levels of violence at regions of origin to generate more displacement, as well as higher levels of violence at destination municipalities and regions to decrease the number of displaced people arriving there. However, these expectations were not matched by our results: intensity of violence at the origin region, and at the destination municipality or region was not statistically significant.

One possible explanation for this lack of significance is that victims of displacement are already being forced to leave their homes, so regional violence would not necessarily be a deciding factor in the impulse to leave. In addition, between 1986 and 2015, almost 35% of displacement occurred within the region of origin and more than 30% occurred within the same municipality. For these displaced people, the origin region is partially or fully equivalent to the region of destination, making the coefficient of these variables insignificant and inconsistent.

In addition, we find there might be two possible explanations for why violence at the destination municipality or region was not significant. First, displaced people might not have the time and luxury to

Table 4
Regression results for annual data (1986–2015).

Dependent variable: displacement (M)	OLS regression				GLM regression			
	Annual data		Aggregate data		Annual data		Aggregate data	
	Basic model	Final model	Basic model	Final model	Basic model	Final model	Basic model	Final model
Constant	0.653	−2.621	0.699	−3.174	−0.192	−0.549	−0.217	−0.625
	(0.135)	(0.276)	(0.089)	(0.173)	(0.101)	(0.140)	(0.055)	(0.078)
Population (P_O)	0.054	0.092	0.078	0.122	0.028	0.050	0.037	0.060
	(0.011)	(0.010)	(0.007)	(0.006)	(0.008)	(0.009)	(0.005)	(0.005)
Population (P_D)	0.150	0.044	0.220	0.108	0.084	0.022	0.109	0.049
	(0.007)	(0.010)	(0.005)	(0.006)	(0.005)	(0.009)	(0.003)	(0.005)
Distance (D)	−0.136	−0.089	−0.218	−0.155	−0.063	−0.038	0.083	−0.056
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.001)	(0.002)
Previous displacement (N_s)		0.170		0.277		0.096		0.114
		(0.005)		(0.004)		(0.005)		(0.003)
Previous displacement (N_d)		0.048		0.051		0.029		0.032
		(0.008)		(0.004)		(0.008)		(0.004)
Violence (V_O)		0.147		0.170		0.088		0.089
		(0.009)		(0.005)		(0.008)		(0.004)
R-squared	0.137	0.339	0.180	0.414				
Adj. R-squared	0.136	0.338	0.180	0.414				
F-statistic	678.1	1026	2433	3831				
N	10,233	10,233	28,700	28,700	10,165	10,165	28,634	28,634

Note: Results were averaged over all different time series.

Table 5
Result of regression with community participation data (2005).

Dependent variable: displacement (M)	OLS regression	GLM regression
Constant	1.1638*** (0.091)	0.1402* (0.068)
Population (P_O)	0.0979*** (0.007)	0.0451*** (0.005)
Population (P_D)	0.0770*** (0.008)	0.0318*** (0.006)
Distance (D)	−0.1083*** (0.004)	−0.0375*** (0.002)
Previous displacement (N_s)	0.1902*** (0.003)	0.0982*** (0.003)
Previous displacement (N_d)	0.0541*** (0.007)	0.0298*** (0.006)
Violence (V_O)	0.1536*** (0.006)	0.0751*** (0.005)
Community participation (C_O)	−0.1229 (0.067)	−0.0796 (0.051)
Community participation (C_D)	0.4309*** (0.095)	0.1605* (0.074)
R-squared	0.400	
Adj. R-squared	0.400	
F-statistic	1456	
N	17,486	17,445

Note: *p < 0.05; **p < 0.01; ***p < 0.001.

comprehensively consider all the possible destinations and their levels of violence. They might just want to seek refuge with someone they know, or where other people from their same area have previously moved to: the closer the better, the sooner the better. Second, between 1986 and 2015 approximately 40% of victims were displaced from the same origin more than once. In this sense, victims of displacement might consider this to be a temporary condition and thus not put so much weight on the levels of violence at their "temporary" destination.

In addition, we found that as the total amount of displacement increased, so did the R-squared of our model. However, over half of the variance of forced displacement remains unexplained by our regressions. Some of the factors that fell outside the scope of this study but might also affect flows of displacement are: the number and quality of government and community institutions that provide aid to migrants, income levels of displaced people, and the specific dominant armed group in each region. Evidence from Syrian (Alhanaee & Csala, 2015) and Colombia (Lozano-Gracia et al., 2010) points towards government and community institutions as one of the significant factors in the choice of destination by displaced people. Additionally, individuals with higher income levels tend to be able to move to farther and safer destinations, while people with lower income tend to have more limited destination choices. Finally, some victims of displacement might want to migrate to locations dominated by armed groups they think might provide protection (Steele, 2009). In addition, as Steele (2009) suggests, some armed groups hold strict control over how displaced people settle in their neighborhoods, which would ultimately affect people's final choice of destination.

Overall, our study highlights and quantifies the effects of population, distance, violence, and, specially, social networks in driving displaced people's choice of destination. These results underscore the regional character of forced displacement, and point towards medium-sized regional centers as crucial sites for attention, investment, and intervention: these cities are close enough to sites of violence, and big enough to be obvious choices for displaced people to flee to. Furthermore, once these medium-sized cities start receiving displaced people, their pull factor will only grow stronger. Planners, government officials, and aid organizations would be remise if they did not focus their attention on these places, allocated resources accordingly, and planned for future arrivals.

In Colombia, as well as in other countries, forced displacement is often seen as a national-level phenomenon: much attention and resources have gone to the country's three largest cities, Bogotá, Medellín, and Cali, which collectively have received almost a million displaced people. However, the lion share of displacement has gone to smaller, regional centers. This, paired with our results points towards a different understanding of forced displacement, one that sees it more as a regional-level issue that should be taken into account in any long-term regional or municipal level planning and resource allocation (Albuja & Ceballos, 2010).

In addition, given the amount of influence social networks have in driving people's choice of destination, planners and aid organizations should proactively engage and collaborate with community organizations, specially victims groups. These grassroots organizations can provide local, on-the-ground knowledge not only about the conditions of displaced people already residing in cities, but also about possible future arrivals based on the conditions at the origin municipalities. With their aid, city officials, planners, and aid organizations, will be able to better understand forced migration and thus formulate more adequate policies and plans, secure resources, and provide services to these vulnerable populations.

5. Conclusion

This research presented an exploratory analysis of some of the factors associated with the choice of destination for victims of forced displacement during the Colombian conflict. We followed and extended the traditional gravity model of migration and included additional variables associated with forced displacement. Our results suggest that violence at the origin municipality is the most important driving factor of displacement. They also suggest that the strongest predictor of displacement volume arriving at a destination is the number of displaced people from the same origin in the previous two years. Distance from origin to destination, and population size at destination are additional key factors in the choice of destination. From this we conclude that forced displacement should be treated as regional-level phenomenon and that planners, city officials, and aid organizations focus their attention on medium-sized regional centers. Because of the importance of social networks in driving people's choice of destination, planners and aid organizations should closely collaborate with grassroots community organizations to adequately allocate resources and plan for new arrivals.

Table 6
Percentage of displaced people by subsets (1986–2015).

Range	Intra-municipality			Inter-municipality		
	31.51%			68.49%		
Gender	Male 48.52%			Female 51.48%		
Age	Young 49.77%		Adult 47.35%		Elderly 2.86%	
Armed group	Guerrillas 39.96%	Paramilitaries 19.89%	Fuerza Pública 0.72%	Bacrim 4.27%	Other 11.19%	Non-identified 27.38%

Declaration of Competing Interest

The authors declare that there is no conflict of interest.

Acknowledgements

This study is part of the Conflict Urbanism: Colombia project by the Center for Spatial Research at Columbia University. Conflict Urbanism: Colombia is Laura Kurgan (principal investigator), Juan Francisco Saldarriaga (research scholar), Dare Brawley, Anjali Singhvi, Patrick Li, Stella, Ioannidou, Mike Howard, Jeevan Farias and Yuan Hua (research assistants). Funding for this project generously provided by the Andrew W. Mellon Foundation. Special thanks to Angelika Rettberg, Associate Professor and Director of the Research Program on Armed Conflict and Peacebuilding, Universidad de Los Andes, Bogotá, Colombia, for facilitating the data and much of the underlying concepts of the conflict.

Funding acknowledgement

Funding for this project generously provided by the Andrew W. Mellon Foundation. The Foundation was not involved in the design of the study, in the collection, analysis and interpretation of the data, or in the writing of the report.

References

Albuja, S., & Ceballos, M. (2010). Urban displacement and migration in Colombia. *Forced Migration Review, 34*, 10–11.
Alhanaee, A., and Csala, D. (2015). Motives to flee: Modeling Syrian refugee crisis. Working paper. [Online] Available at https://www.researchgate.net/publication/299049673 [Accessed Apr 28, 2017] DOI:10.13140/RG.2.1.2986.1524.
Bartel, Ann P. (1989). Where Do the New U.S. Immigrants Live? *Journal of Labor Economics, 7*(4), 371–391.
Borjas, George J. (1987). Self-Selection and the Earnings of Immigrants. *The American Economic Review, 77*(4), 531–553 (23 pages).
Borjas, G. J. (1989). Economic Theory and International Migration. *International Migration Review, 23*(3), 457–485.
Crisp, J., Morris, T., & Refstie, H. (2012). Displacement in urban areas: New challenges, new partnerships. *Disasters, 36*(S1), S23–S42.
Dahl, M., & Sorenson, O. (2010). The migration of technical workers. *Journal of Urban Economics, 67*, 33–45.
Darchen, S., & Tremblay, D. G. (2010). What attracts and retains knowledge workers/students: The quality of place or career opportunities? The cases of Montreal and Ottawa. *Cities, 27*, 225–233.
Departamento Administrativo Nacional de Estadísticas (2005). Estimación y proyección de población nacional, departamental y municipal total por área 1985–2020. [Online] Available at https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/proyecciones-de-poblacion, Accessed date: June 2019.
Engel, S., & Ibáñez, A. (2007). Displacement Due to Violence in Colombia: A Household-Level Analysis. *Economic Development and Cultural Change, 55*(2), 335–365.
Frenkel, A., Bendit, E., & Kaplan, S. (2013). Residential location choice of knowledge workers: The role of amenities, workplace and lifestyle. *Cities, 35*, 33–41.
Gossaín, J. (2017). La espantosa historia de las falsas víctimas del conflicto. *El Tiempo*. [Online] Available at http://www.eltiempo.com/justicia/investigacion/la-espantosa-historia-de-las-falsas-victimas-del-conflicto-96674, Accessed date: August 2019.
Greenwood, Michael J. (1975). Research on Internal Migration in the United States: A Survey. *Journal of Economic Literature, 13*(2), 397–433.
Greenwood, M. (2005). Modeling migration. *Encyclopedia of social measurement*. *2. Encyclopedia of social measurement* (pp. 725–734).
Ibáñez, A. (2009). Forced displacement in Colombia: Magnitude and causes. *The Economics of Peace and Security Journal, 4*(1), 48–54.
Ibáñez, A., & Moya, A. (2006). The impact of intra-state conflict on economic welfare and consumption smoothing: Empirical evidence for the displaced population in Colombia. *HiCN working papers 23. Households in conflict network*. [Online] Available at https://ideas.repec.org/p/hic/wpaper/23.html, Accessed date: 28 May 2018.
Ibáñez, A., & Velásquez, A. (2009). Identifying victims of civil conflicts: An evaluation of forced displaced households in Colombia. *Journal of Peace Research, 46*(3), 431–451.
Ibáñez, A., & Vélez, C. (2008). Civil conflict and forced migration: The micro determinants and welfare losses of displacement in Colombia. *World Development, 36*(4), 659–676.
Kalyvas, S. N. (2006). *The Logic of Violence in Civil War.* Cambridge; New York: Cambridge University Press.
Kaplan, S., Grünwald, L., & Hirte, G. (2016). The effect of social networks and norms on the inter-regional migration intentions of knowledge-workers: The case of Saxony, Germany. *Cities, 55*, 61–69.
Karemera, D., Oguledo, V., & Davis, B. (2000). A gravity model analysis of international migration to North America. *Applied Economics, 32*(13), 1745–1755.
Lawton, P., Murphy, E., & Redmond, D. (2013). Residential preferences of the 'creative class'? *Cities, 31*, 47–56.
Lozano-Gracia, N., Piras, G., & Ibáñez, A. (2010). The journey to safety: Conflict-driven migration flows in Colombia. *International Regional Science Review, 32*(2), 157–180.
Office of the United Nations High Commission for Refugees (2019). Global trends: Forced displacement in 2018. [Online] Available at https://www.unhcr.org/globaltrends2018/, Accessed date: 20 June 2019.
Poor, J., Alimi, O., Cameron, M., & Maré, D. (2016). The gravity model of migration: The successful comeback of an ageing superstar in regional science. *IZA discussion paper no. 10329*. [Online] Available at http://ftp.iza.org/dp10329.pdf, Accessed date: 28 May 2018.
Red Nacional de Información (2016). *Base de Datos Víctimas.* (Unpublished raw data).
Republic of Turkey, Ministry of Interior (2019). Distribution of Syrian refugees in the scope of temporary protection by province. [Online] Available at https://www.goc.gov.tr/icerik6/temporary-protection_915_1024_4748_icerik, Accessed date: 20 June 2019.
Schmeidl, S. (1997). Exploring the causes of forced migration: A pooled time-series analysis, 1971–1990. *Social Science Quarterly, 78*(2), 284–308.
Stanley, W. (1987). Economic migrants or refugees from violence? A time-series analysis of Salvadoran migration to the United States. *Latin American Research Review, 22*(1), 132–154.
Steele, A. (2009). Seeking safety: Avoiding displacement and choosing destinations in civil wars. *Journal of Peace Research, 46*(3), 419–430.
Zipf, G. (1946). The P_1 P_2/D hypothesis: On the intercity movement of persons. *American Sociological Review, 11*(6), 677–686.