Time sequence of forced displacement into neighbouring countries

Mattia Di Salvo, Alexandre Ounnas & Doina Postica

Centre for European Policy Studies

February 2022

"This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nº 882986".

Deliverable Factsheet	
Title and number	*Time sequence of forced displacement into neighbouring countries (D.3.4)*
Work Package	WP3
Submission date	24/February/2022
Authors	Mattia Di Salvo (CEPS), Alexandre Ounnas (CEPS) & Doina Postica (CEPS)
Contributor(s)	N/A
Reviewers	Alireza Jahani (BUL) & Fouad Shammary (MTU)
Dissemination level	PU (Public)
Deliverable type	R (Report)

Version Log

Issue Date	Version nº	Author	Change
31/01/2022	v0.1	Mattia Di Salvo, Alexandre Ounnas & Doina Postica	First version sent for review.
15/02/2022	v0.2	Mattia Di Salvo, Alexandre Ounnas & Doina Postica	Second version after comments
22/02/2022	V0.3	Mattia Di Salvo, Alexandre Ounnas & Doina Postica	Final version
22/02/2022	V0.3	Cristina Blasi	Coordinator final review.

Executive summary

The International Organization for Migration (IOM) defines forced migration or forced displacement as migratory movement induced by several factors such as force, compulsion, or coercion. The United Nations High Commissioner for Refugees (UNHCR) reports that the number of displaced individuals almost doubled over the last decade, with **around 40% of these individuals being compelled to cross borders.**[1] These refugees or asylum seekers **settle primarily in neighbouring countries**, which are usually developing countries with limited resources and often precarious political situations.

This report aims to better understand cross-border displacement by analysing the **time sequence from events potentially inducing displacement to migration into the neighbouring country**. This time sequence is of particular interest to many stakeholders, as better knowledge of the timing can help decision makers anticipate and plan possible actions to support host countries. For instance, it can help the planning of humanitarian corridors and resettlement pledges, as well as the support to international organisations providing shelter and assistance in refugees camps. Furthermore, timing on initial cross-border displacement can also inform about possible developments of mixed-migration along the migratory routes to Europe.

The data on the number of people displaced across borders is obtained from the UNCHR Operational Data Portal (ODP). The analysis covers seven origin countries and twenty-one neighbouring countries at a monthly frequency over several years up to the end of 2020. Two types of events are studied as possible factors in origin countries leading to displacement. Firstly, violent events due to conflict are extracted from the Georeferenced Event Dataset from the Uppsala Conflict Data Program (UCDP). Secondly, mass disasters such as floods or earthquakes, are retrieved from the Emergency Events Database (EM-DAT) of the Centre for Research on the Epidemiology of Disasters (CRED).

Using this longitudinal dataset, the analysis applies a finite distributed lag (FDL) model with fixed effects and carefully selects the number of lagged events relevant

[1] This figure refers to the total number of displaced individuals at the end of 2020.

to explain displacement into neighbouring countries. The results indicate that **conflict events can have a positive long and lasting effect on displacement and significant effects are found five months after conflict events took place**. Moreover, the response of **cross-border displacement to conflict is hump-shaped**, with the largest effect found one month after the event. The results for **mass disasters** are more mixed: although estimated coefficients are larger than those for conflict, the only positive and significant (at the 10% level) effects are found **one and four months after the event**.

Therefore, findings provide evidence of the stronger impact of conflict compared to mass disasters on displacement into neighbouring countries. Their effects are persistent and can last for several months after the occurrence of the events themselves. However, it is important to keep in mind that mass disasters are considerably less frequent than conflict events. Moreover, disasters can indirectly increase cross-border displacement when their effects exacerbate instability and fuel conflicts in origin countries.

Using a multi-variate approach with data covering different origin and destination countries for several years, this report is important to provide robust evidence at a macro-level that can inform the small-scale model within the EMT. For instance, information on the time sequence of displacement can help calibrate model's parameters for simulation of future trends. Similarly, the use of different set of regressors provide further evidence on the robustness of results, as well as of their relevance in driving cross-border displacement, e.g. differentiating between conflict and mass disasters, or between mass disasters and significant changes in weather conditions.

Keywords: Forced displacement, cross-border, time sequence, neighbouring countries, conflict and mass disasters.

Table of Contents

List of Tables

List of Figures

Abbreviations

AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
CRED	Centre for Research on the Epidemiology of Disasters
EM-DAT	Emergency Events Database
EMT	EUMigraTool
EU	European Union
FAO	Food and Agriculture Organization
FDL	Finite Distribute Lag
GED	Georeferenced Event Dataset
GDP	Gross Domestic Product
IDMC	Internal Displacement Monitoring Centre
IOM	International Organization for Migration
ODP	Operational Data Portal
PPP	Purchasing Power Parity
REIGN	Rulers, Election, and Irregular Governance
SA	Seasonal Adjustment
SPI	Standardized Precipitation Index
UCDP	Uppsala Conflict Data Program
UNHCR	United Nations High Commissioner for Refugees
WDI	World Development Indicators

1 Introduction

Over the past decade, the European Union has recorded several increases in the number of people arriving from different regions to seek protection. However, despite the important role of some EU countries in granting asylum and providing resettlement, EU and developed countries host a very limited share of the global number of people displaced abroad. In 2020, the United Nations High Commissioner for Refugees (UNHCR, 2021) reported that 86% of refugees and asylum seekers were hosted in developing countries and, most importantly, **73% were in neighbouring countries**. These figures clearly show how forced displacement across borders has a strong regional component, with countries in the immediate proximity representing the first natural place to reach for forcibly displaced people.

Hence, given their key role as host countries, it is important to understand how destabilising events in origin countries, such as conflict and natural disasters, influence displacement into neighbouring countries, and to what extent these effects are prolonged over time. Evidence on the **time sequence from incidents to cross-border displacement** can help inform decision makers about possible actions in support of host countries and international organisations working on the ground. It can also help them to anticipate and plan humanitarian corridors and resettlement pledges with a proactive rather than reactive approach. This is even more important when considering that host neighbouring countries are most often developing countries with limited resources, and can themselves face precarious political stability, conflicts and generalised violence.

The aim of this report is to understand which and how many time-lags of changes in events capturing conflict or mass disasters in origin countries are significant in explaining changes of displaced people in neighbouring countries. This is key to provide **early warnings and support** to neighbouring countries and humanitarian organisations, and thus planning of resource allocation for refugees camps. In this respects, results on the time-sequence of displacement due to conflict and disasters provide robust evidence at a macro-level that can better inform the multiscale simulation within the EMT.

Furthermore, this evidence can help **foresee possible developments in mixed-migration flows to the EU**, as neighbouring host countries often represent the starting points of journeys along the migration routes. In this respect, this analysis complements findings from Deliverable 3.1 on the effects of disruption and growing instability in transit countries on mixed-migration flows to the EU.

The next section reviews the concept of forcibly displaced populations and the impact that conflict and disasters – for instance natural disasters - have on displacement. Section 3 reports on data sources and explains the types of data used, while section 4 lays down the methodology used for the empirical exercise. Section 5 present the results and section 6 concludes.

2 Literature review

Despite lacking international legal recognition, the concept of forced displacement has been widely used to describe the movements of refugees, asylum seekers, internally displaced persons and even victims of trafficking.[2] Overall, forced displacement is usually induced by different factors such as force, compulsion or coercion.[3]

According to the UNHCR, the number of internal and cross-border forced displacement nearly doubled in the last decade. By the end of 2020, 26.4 million refugees and 4.1 million asylum-seekers accounted for 39% of the total number of forcibly displaced people.[4]

In the literature, **conflict- and disaster-induced forced displacement** is investigated under different specifications, where conflict-induced displacement is caused by humans, whereas disaster-induced displacement originates from natural causes. Moreover, it is important to keep in mind that these types of events can affect each other, as conflicts may arise in periods of adverse climate events or human activity may cause natural disasters. Additionally, while the literature analysing the main drivers of migration is vast, very little is available on the timing and dynamic effects of conflict and disasters on cross-border forced displacement.

The analysis of conflict- and disaster-induced displacement is therefore particularly important since many of the **countries facing adverse effects of climate change often face political instability and conflicts as well**. The Internal Displacement Monitoring Centre (IDMC) estimates that 95% of internal displacement in 2020 was caused by conflicts in countries vulnerable to climate change. Moreover, even though the relationship between internal and cross-border

[2] According to the UNHCR, refugees include groups of people recognised under the 1951 Refugee Convention, its 1967 Protocol, and the 1969 Organization of African Unity (OAU) Convention. Since 2007, the term has included "groups of persons who are outside their country or territory of origin and who face protection risks similar to those of refugees, but for whom refugee status has, for practical or other reasons, not been ascertained" (UNHCR, 2017). Asylum seekers, on the other hand, include people who have not received a final decision on their application for asylum or refugee status. For detailed definitions see UNHCR Statistical Online Population Database: Sources, Methods and Data Considerations.

[3] Forced migration or displacement data (migrationdataportal.org)

[4] In 2020, according to the UNHCR (2021), 68% of refugees were from five countries, among which the Syrian Arab Republic (6.7 million), Venezuela (4 million) and South Sudan (2.2 million).

displacement is not well understood yet, **countries with a high number of internally displaced people also become major origin countries of asylum-seekers and refugees** (IDMC, 2019). Since asylum seekers and refugees are clustered in specific regions characterised by overall instability, the situation in neighbouring countries must be monitored closely in order to avoid further deterioration in the regions concerned.

The literature studying the relationship between climate and conflict, as well as their impact on migration, has been developing over the past two decades. In a study of patterns of international migration in 45 Sub-Saharan African countries, Naudé (2010) finds that **violent conflict and gross domestic product (GDP) growth differentials have the largest impact on international migration in the region**. He concludes that international migration from Sub-Saharan Africa is both an adapting and mitigating strategy in the face of conflict and economic stagnation. Using panel data on intraregional migration in Sub-Saharan Africa, Ruyssen et al. (2014) find that the occurrence of conflict in one country encourages migration **towards countries that are relatively freer** and that the creation of networks has a positive effect on migration by lowering the psychological costs. They also suggest that distance is an important factor, because of the role played by transport and communication costs. Moreover, Moore et al. (2004) find evidence that past movements increase current ones either through the creation of networks that reduce the cost of moving for subsequent persons or by **increasing the cost of staying in violent contexts**.

While studies on the direct impact of climate change on migration flows are scarce, a strand of the literature highlights a potential **link between climate change and raising conflicts**. Using cross-border refugee flows between 1989 and 2014, Abel et al. (2017) confirm that displacement from countries experiencing high conflict intensity are higher than those with low conflict. Second, they report that **conflict outbreaks are more probable in countries with scarce freshwater resources and a lower level of rainfall**. In a subsequent study, Abel et al. (2019) exploit bilateral data on asylum applications for 157 countries over the period 2006–2015, finding evidence in support of the link between climate and conflicts but only for the period 2010-2012, where the dynamics of asylum seekers were dominated by the Arab spring, as well as by war episodes in Sub-Saharan Africa.

Abel et al. conclude that the adverse effects of climate change on asylum seeker dynamics are **more likely to emerge in countries undergoing political transformation**, particularly in the form of population discontent with current governments. A meta-analysis of 60 quantitative studies further confirm that the **risk of conflict increases with deviations from normal precipitation and mild temperature** (Hsiang S. M., 2013). Finally, Burke et al. (2015) estimate the cumulative effects of climate variables on conflict. They find that, on average, interpersonal conflict rises by 1.2% and intergroup conflict by 4.5% for each standard deviation change in climate toward more adverse conditions. In a recent study, Jahani et al. (2021) implement a multiscale (macro- and micro-scale) simulation approach to model forced displacement for the case of South Sudan. In the microscale model, the authors define three types of routes (i.e. drive, walk and river) and include weather conditions to determine realistic agent movements (e.g. changes in road accessibility due to flooding). While the multiscale model with weather conditions has a significantly longer execution time than the others and thus need to be improved, the validation error is rather similar to that of the single-scale model, indicating that the choice of destination and thus of crossing point can be affected by weather conditions.

3 Data

As reported in the literature, forcibly displaced populations because of e.g. persecution, conflict or generalised violence encompass a wide range of population groups. These include refugees and asylum seekers when referring to people displaced across borders, or internally displaced people when referring to, indeed, displacement within borders. In 2020, around 30.5 million people were forcibly displaced across borders, **representing about 40% of the total number of people forcibly displaced worldwide** (UNHCR, 2021). Hence, given the focus of the analysis on the time sequence of displacement from country of origin to neighbouring countries, statistics on the number of forcibly displaced people across borders are collected.

The **UNHCR Operational Data Portal (ODP)** provides detailed and up-to-date information on several countries and situations of concern monitored. The portal was created in 2011 to fulfil the UNHCR's institutional responsibility to provide an information and data-sharing platform. The ODP aims to facilitate the coordination of major refugee emergencies by mean of independent 'situation views', such as those in Syria and the Central African Republic.[5] As opposed to standard yearly statistics on refugees and asylum seekers, figures collected under the ODP are published at a significantly higher frequency, which can be as often as weekly, to allow for more in-depth focus and analysis of the number of people displaced abroad.

This **high frequency** of the data is a key aspect in the analysis of the time sequence of displacement driven by conflict and disasters in origin countries. **Event-based data** on conflict and disasters, in fact, report a start and end date that permit figures to be easily aggregated at different frequency levels, thus allowing better identification of their effects on changes in the number of displaced people.

Data collected on forcibly displaced populations cover seven origin countries and twenty-one neighbouring countries (i.e., situations monitored under the ODP).[6] Depending on each situation's characteristics, the time span and frequency of

---

[5] https://data2.unhcr.org/en/about/

[6] For Nigeria, the breakdown by neighbouring country is not available, so the total displacement in Niger, Chad and Cameroon is used. See Table 4 in Appendix for details.

observations can differ. For instance, while information on the situation in Syria goes back to 2012, other series can start as late as 2018.[7] However, the figures reported go up to the current date, thus permitting recent developments to be taken into account in the analysis. In this case, the analysis uses December 2020 as the cut-off date, due to data limitations on conflicts and disasters in 2021. In terms of time frequency, several series are originally reported monthly while others can have a fortnightly or even weekly frequency. All series have therefore been harmonised to a **monthly frequency** by taking the latest figure available for each month on the total number of displaced people present in a neighbouring country. When breaks in series were found over a few months, we used linear interpolation between known monthly observations to estimate the missing values.

Event-based datasets are used to capture growing instability, violence and mass disasters in countries of origin that would eventually lead to displacement into neighbouring countries. **The Georeferenced Event Dataset from the Uppsala Conflict Data Program (UCDP)** records violent events due to conflict[8], while the **Emergency Events Database (EM-DAT) from the Centre for Research on the Epidemiology of Disasters (CRED)** records mass disasters that have impacted the population. Mass disasters recorded include mostly natural disasters related to climate (e.g. floods, droughts) but also other types of disasters impacting the population at a large scale (e.g. earthquakes). While violent events due to conflict usually begin and conclude on the same day, mass disasters can have a longer development depending on their nature, such as a period of drought. In these cases, events are counted not only in their starting month, but in all the months of their duration.[9]

As a robustness check, we use alternative indicators to event-based data to capture the level of political stability and changes in weather conditions. From the **Rulers, Elections, and Irregular Governance (REIGN)** dataset, we retrieve an index on

[7] See Table 4 in the Appendix for the time coverage.

[8] More specifically, violent events deriving from conflict are events that recorded casualties.

[9] Since these are event-based datasets, start and end dates are provided, allowing the duration in days of an event to be computed. The start and end dates are complete for most observations. In a few cases, however, only year and month are available for the end date, in which cases the first day of the month is imputed.

political violence,[10] as well as the Standardized Precipitation Index (SPI). From the **Food and Agriculture Organization (FAO)**, instead, we retrieve temperature change. All three indicators are expressed as deviations from the (long-term) average computed by country and month, where zero values reflect no change from the baseline.[11] In this respect, therefore, these indices capture periods characterised by significant deviations from the 'standard' situation. Moreover, the precipitation and temperature indices focus on the changes in weather conditions rather than on mass disasters.

Finally, other variables from the **World Development Indicators (WDI) of the World Bank** are added to the model to control for socioeconomic dimensions in the origin countries. As explained in Deliverable 3.1 of the project, the main countries of interest under analysis suffer from serious data limitations for several common indicators. Therefore , the main specification includes population, GDP per capita in purchasing power parity (PPP) (constant 2017 international $), and youth unemployment rate (15-24 years).[12] To harmonise the time frequency from yearly to monthly, temporal disaggregation is applied, which is the process of deriving high frequency time series from low frequency ones (also known as target series) and is widely used in official national statistics (Sax, 2013).[13] Finally, in order to also cover 2020 for which data are available for the two main variables of interest (number of displaced people and disruptive events), 2020 figures for control variables have been extrapolated using the countries' historical trends.

[10] This is based on information from the UCDP as well as the START Global Terrorism Database

[11] The political violence and precipitation indices can be interpreted as a z-score on the historically expected values, while temperature change is expressed in °C with respect to the baseline period 1951-1980.

[12] In fact, population, GDP per capita and unemployment rate were the statistics most widely available for all countries. Nevertheless, missing information also concerned these variables in some cases: for GDP per capita, series have been estimated using the average of the region to which the country belongs (e.g. Middle East for Syria). The countries concerned are Syria and South Sudan in the main specification, as well as Venezuela when included in the robustness check.

[13] The Denton-Cholette mathematical-based method was used to interpolate or distribute monthly data, using the sum and the last conversion type for flow and stock variables respectively. See Deliverable 3.1 for more details.

4 Methodology

For the purpose of this study, we use panel data with combinations of country of origin and country of destination. The main advantage of panel data for cross-border forced displacement is that we can model the relation between the occurrence of disruptive events and people's displacement over time. More specifically, it allows us to test at **which lags the occurrence of an event in origin countries can explain displacement into neighbouring countries**.

The baseline specification of our model is:

$$\Delta \ln I_{ijt} = \beta_0 + \sum_{k=0}^{p} \beta_{con,k} \Delta CONFL_{i,t-k} + \sum_{k=0}^{p} \beta_{dis,k} \Delta DISAS_{i,t-k} + \delta X' + \gamma_{ij} + \mu_y + e_{it} \quad (1)$$

Where I_{ijt} denotes the stock of people displaced from country *i* to country *j* at time *t*. $CONFL_{it}$ and $DISAS_{it}$ denote the number of violent events due to conflict and of mass disasters in country *i* at time *t*. These events can influence displacement contemporaneously and **up to *p* lags**, where *p* is the same for both variables. $\mathbf{X}'$ is a vector with the additional explanatory variables, $\boldsymbol{\delta}$ is the vector of corresponding coefficients.

In the vector of explanatory variables $\boldsymbol{X}' = [\ln \Delta DGDP_{it}, \ln \Delta GDP_{it}^2, \ln \Delta POP_{it}, \Delta UNRT_{it}]$, GDP_{it} is the GDP per capita in PPP, POP_{it} is population and $UNRT_{it}$ is the unemployment rate of people between 15 and 24 years old in country *i* at time *t*. Finally, β_0 is the constant, γ_{ij} are origin/destination countries fixed effects and μ_y are year indicator variables.

We estimate a model in which the two main explanatory variables of interest, the number of violent events due to conflict and mass disasters, are included in the estimated equation with *p* lags. Thus, we estimate a **finite distributed lag (FDL) model** recognising that events can have **an immediate effect on cross-border displacement and also a delayed and persistent impact** due to the time required to decide to leave the country and actually do so.

To give an idea about the interpretation of the estimates in a FDL model with a temporary change in the explanatory variables, let us assume we have a model with two lags:

$$y_t = \alpha_0 + \beta_1 z_t + \beta_2 z_{t-1} + \beta_3 z_{t-2} + e_t$$

Here, the coefficient β_1 of the variable at time t is the immediate change in the number of people displaced due to a one unit increase in the explanatory variable at time t. This coefficient is usually called the **impact propensity** or **impact multiplier**. Similarly, the coefficient on the first lag β_2 would be the change in the dependent variable one period (t+1) after the temporary change in the explanatory variable, and so on. In the third period, it is assumed that the explanatory variable has no effect on the dependent one.

When we graph the coefficient of the lagged variables as a function of the number of lags, we obtain the **lag distribution,** which represents the dynamic effect of a temporary increase in the explanatory variable on the number of people displaced (see Figure 2 in the Results). The lag distribution shows the pattern and at which lag there is the largest effect of the explanatory variable on the dependent one (Wooldridge J. M., 2015).

Our specification of the FDL model with panel data is estimated with standard **ordinary least squares in first differences** to account for autocorrelation, and includes time dummies to control for additional yearly effects not captured by our specification.

The absence of serial correlation between first differenced error terms is confirmed by the Wooldridge test (Wooldridge J. M., 2002) (Drukker, 2003). The test is performed on the residuals from the entire dataset and on the residuals for each bilateral pair of origin country and country of destination.

Finally, we check for possible outliers in the first-difference of our main variable. While positive outliers can in fact be explained by significant spikes in, for instance, conflict at origin, negative ones can mostly be due to breaks or adjustments in data collection. We therefore decide to drop the negative changes overcoming the 0.25 threshold in the first difference of the log of displaced people, our main variable. In the end, only 0.5% of our observations are dropped.[14]

[14] Further test on both the entire sample as well as a sample dropping negative and positive outliers show similar results, with lags of violent events due to conflict being always significant and of mass disasters only for selected lags. When also positive outliers are dropped, coefficient of violent events are overall smaller. Moreover, we perform a more formal outlier correction using the software JDEMETRA+ in section 5.3 as an additional robustness check on our results.

5 Results

5.1 Lag order selection

Our interest is to determine how many lags (i.e., months) are significant in explaining cross -border displacement from a country of origin to a country of destination. Evidence on the dynamics of forced displacement when, for instance, a conflict intensifies is important to better formulate actions in support to people in need: for instance, increasing the capacity of refugees camps in neighbouring countries. In this respect, these findings can help improve the accuracy of predictions under the small-scale model of the EMT both by providing estimates on changes in number of displaced people as well as the time-sequence of these effects across different months.

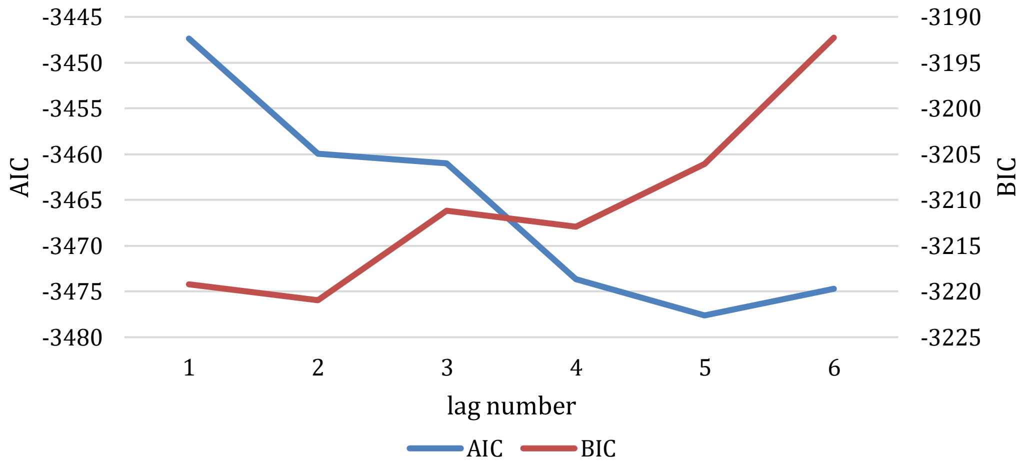
There is no a priori way to determine the appropriate lag length, and the selection of the number of lags to include is usually determined empirically. There are several methods in econometrics theory that allow to gain information about the appropriate lag length, among which the F-test and Akaike and Bayesian Information Criteria (AIC and BIC). One possible drawback of lag selection based exclusively on the F-test is that it can select models that are too large, thus we determine the appropriate lag length upon the results provided by the AIC and BIC, while controlling for the F-test and the adjusted-R^2.

These information criteria trade off goodness of fit (as measured by the log likelihood) with over-parametrisation of the model. They therefore allow for the selection of a parsimonious lag order specification that nonetheless fits the data well enough. When comparing the AIC and BIC, the latter penalises more for additional parameters which usually leads to the selection of simpler specifications. In this sense, the BIC is more conservative than the AIC (Enders, 2014).

In practice, we need to select a maximum lag order to consider, $\bar{p}$. We then estimate model (1) for each possible value of $p \in [0;\ \bar{p}]$ and record the AIC and BIC. **The minimum values of the criteria indicate the possible best specifications.**

To select the maximum number of lags, $\bar{p}$, the main restriction lies in the fact that the larger $\bar{p}$ is, the more data points are lost at the start of the sample to generate the lag values of the explanatory variables. Given that the sample size for some countries is relatively small, we set the maximum number of lags to be tested to six[15]. Figure 1 displays the results and indicates that the AIC would select five lags of the explanatory variables, while the BIC reaches its minimum in lag two and starts increasing from lag four.

Figure 1: Information criteria for a maximum number of lags set to six.



To further help us determine our preferred number of lags, we estimate model (1) for two to six lags, and check for the statistical significance of the estimated parameters. The full set of results is displayed in Table 5 in the Appendix. The results indicate that parameters of the two variables of interest can be statistically significant **up to lag five,** which corresponds to the number of lags selected by the AIC. We therefore select this number for our preferred specification.

5.2 Estimated impact on displacement

Table 1 presents the estimated parameters for violent events due to conflict and mass disasters. Given that our dependent variable is specified in log difference, the coefficients report semi-elasticities.

[15] In Figure 3 in the Appendix, we present results for a maximum number of lags set to 12. The results are then slightly different as the AIC would reach its minimum in the 11th lags of conflict events mass disasters but with the BIC growing steadily from lag one. However, including 12 lags can lead us to drop up to a third of the sample for some countries (e.g., the series for the Democratic Republic of the Congo only starts in 2018).

Starting with conflict, we see that all the coefficients are statistically significant at the 1% level but for the fifth lag (10% level). The estimated parameters imply that **one additional conflict event in month *t-1* increases the current number of people displaced by 0.02%,** holding all others variables constant. While this effect can appear to be small, it is important to keep in mind that an average of 140 events per month are recorded in the countries of origin in our sample[16]. Considering an increase in the number of conflict events from 0 to this average in month *t-1*, for instance, the number of displaced people at time *t* would increase, on average, by 2.67%.

Table 1 Conflict and Disasters – five lags specification

Lag number	*Conflict*	*Disasters*
0	0.00015***	-0.00239
	(0.00005)	(0.00378)
1	0.00019***	0.00654*
	(0.00005)	(0.00367)
2	0.00016***	0.00023
	(0.00005)	(0.00420)
3	0.00010***	0.00279
	(0.00002)	(0.00250)
4	0.00013***	0.00493*
	(0.00003)	(0.00281)
5	0.00008*	-0.00218
	(0.00004)	(0.00231)
Observations	1,688	
R-squared	0.211	
Adjusted R-squared	0.186	

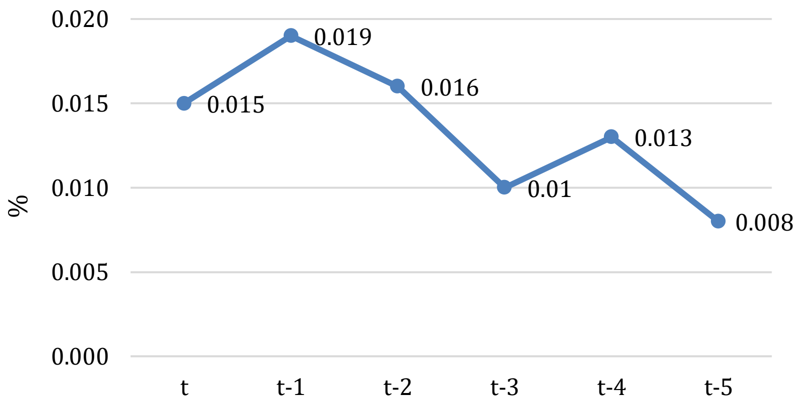
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Moreover, our **dynamic specification implies that an event during month *t* has an effect on cross-border displacement up to month *t+5***. In other words, one event in month *t* leads to a total increase of approximately 0.08% five months after the event. Finally, the **lag distribution** suggests **a hump-shaped response of**

[16] Furthermore, only 1% of observations record no conflict events in a given month.

displacement to the conflict event, with the highest coefficients obtained at lag one (Figure 2).

Figure 2 Lag distribution conflict– percentage change displaced people



With respect to **mass disasters**, we note that estimated coefficients tend to be greater than conflict events ones, **but are significant only for lag one and four at the 10% level**. According to this estimate, one additional mass disaster in month *t-1 (t-4)* leads to a 0.65% (0.49%) increase in the current level of displaced people.

Finally, as regards the control variables included in the model, population and unemployment rate are always positive but not significant across all lags tested, while GDP per capita and its squared term are usually both significant, with a positive and negative sign respectively (see Table 5 in the Appendix). This reflects the reversed U-shaped relationship between economic development and emigration, already assessed in the literature and shown in the results of Deliverable 3.1 of this Task.

Overall, these results suggest that **violence due to conflict plays a more important role in explaining displacement to neighbouring countries compared to mass disasters**. Furthermore, the effects of conflict are persistent and can last for several months after the occurrence of the events themselves. However, it is important to keep in mind that mass disasters are considerably less frequent than conflict events, but are found to have a larger impact when coefficients are significant. Finally, disasters can also be the cause of escalating tensions and conflicts within countries characterised by precarious stability and, therefore, indirectly cause displacement into neighbouring countries (see Literature).

5.3 Robustness checks

5.3.1 Monthly indices

As alternative specification, we use monthly indices to capture levels of political violence, precipitation and temperature vis-à-vis the long-term average of each country (and month when referring to weather conditions). Indices for political violence and precipitation are z-scores, while temperature change is expressed in °C. Using precipitation and temperature indices, this specification tests for the impact of changes in weather conditions on displacement rather than of mass disasters, as done in the main specification. This further test is therefore important to disentangle possible diverging effects of different dimensions on displacement and, therefore, to better inform the modelling of the EMT.

As with the main specification using event-based data, we test for five lags for each of the three indicators: coefficients for the lags of the index on political violence are found to be significant (between 5% and 10% statistical level), while none is for precipitation, and only in period *t* for temperature change (Table 2).

Table 2 Political violence, temp. and precip. indexes indeces – five lags specification

Lag number	*Political violence*	*Temperature*	*Precipitation*
0	0.00173**	0.00565*	-0.00610
	(0.00079)	(0.00324)	(0.00829)
1	0.00364*	0.00432	0.00934
	(0.00194)	(0.00393)	(0.00951)
2	0.00171	0.00521	-0.00347
	(0.00379)	(0.00351)	(0.01849)
3	0.00181*	-0.00033	0.02250
	(0.00098)	(0.00294)	(0.02309)
4	0.00194*	0.00343	-0.01492
	(0.00113)	(0.00316)	(0.01372)
5	0.00256**	0.00373	0.01172
	(0.00127)	(0.00258)	(0.01190)
Observations		1531	
R-squared		0.183	

Adjusted R-squared	0.154

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

However, since these indices capture variations from the long-term average of each country, changes from one month to another might not capture extended periods with significantly worse conditions. We therefore try to test controlling for 'exceptional' months vis-à-vis the average using different thresholds for the three variables: 90% and 95% significance level for the political violence and precipitation indexes (i.e. z-score of at least 1.645 and 1.96), and a change of 1.75 and 2 °C for temperature. Afterwards, we count how many of the previous five months were 'exceptional' in the country of origin.[17] More specifically, we are interested in months that have values above the threshold for political violence and temperature (i.e., periods with much higher levels of political violence and temperatures far above average), and those that are below the threshold for precipitation to capture periods of drought (i.e. periods with much lower precipitation than average).

Results suggest that political violence has a more significant impact on displacement into neighbouring countries than changes in weather conditions. Coefficients for the running count of exceptionally violent months in the period *t-1 to t-5* are positive and significant at 1% (Table 3). No significant results are reported for periods with months that are significantly warmer than average, while significant and negative coefficients are found for periods with precipitation significantly lower than average (even if with a lower statistical significance and magnitude compared to political violence). This negative relationship might indicate that periods with significantly low levels of precipitation hamper possibilities for people to cross borders, because, for instance, of harsh weather conditions along the path to neighbouring countries.

[17] In other words, these variables range from a minimum of zero (for no exceptional month recorded in the previous five months) to five (all five previous months were exceptional).

Table 3 Political violence, temp. and precip. indexes – running count of exceptional months between t-1 and t-5

	Threshold 90% pol. violence, precip 1.75 °C temp	Threshold 95% pol. violence, precip 2 °C temp
Political Violence *(above threshold)*	0.03165*** (0.01003)	0.04792*** (0.01546)
Temperature *(above threshold)*	0.00097 (0.00406)	0.00547 (0.00416)
Precipitation *(below threshold)*	-0.01464* (0.00802)	-0.01871** (0.00877)
Observations	1536	1536
R-squared	0.179	0.184
Adjusted R-squared	0.159	0.163

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

These findings can be further assessed within the modelling under EMT by using information on exact location of precipitations, which allows to differentiate between areas that are along the path to neighbouring countries from those that are densely populated. This is important to take into account possible diverging and second-level effects identified in the literature. Burke et al. (2015) , for instance, find that each standard deviation change in climate toward more adverse conditions increases intergroup conflict by 4.5%, which thus in turn can lead to further displacement. Similarly, conducting a meta-analysis of 60 studies, Hsiang (2013) finds that the risk of conflict increases with deviations from normal precipitation and mild temperature.

5.3.2 Additional tests

While our main sample covers geographic areas close to the EU and thus more relevant in terms of mixed-migration flows, we decide to include also Venezuela in an additional specification. The country is, in fact, among the top origin countries worldwide in terms of displaced population, with, as usual, neighboring countries being the major recipients. Specifically, we include information on the number of Venezuelans displaced in Brazil, Colombia, Ecuador and Peru, with September

2018 as start date.[18] Results are very consistent with those of the main specification, both in terms of coefficients and significance of lags of both conflict and disasters (see Table 6 in the Appendix).

Using our main sample, we test another specification including the auto-regressive factor of the dependent variable. The AR of the first order is significant at 1% level, implying that changes in the number of displaced people at time t are affected by those at t-1. Overall results hold with respect to the main specification: lags of violent events due to conflict are in fact significant up to the fourth lag, with the first two reporting the highest coefficients, while only the first and fourth lags of mass disasters are significant at 10% level (see Table 7 in the Appendix).

Finally, since statistics on the number of displaced population can suffer of different limitations, especially due to the challenging conditions for data collection, we decide to check for possible seasonality in the data and correct for eventual outliers. As explained in the Methodology, we used a 0.25 as threshold to drop negative outliers in our main specification, which accounted for only 0.5% of our sample. In this specification, instead, we used the software JDemetra+[19], which has been developed by Eurostat, the National Bank of Belgium, and the Bundesbank to provide an easy-to-use tool for the seasonal adjustment (SA) of any economic series. JDemetra+ offers two methods to perform the SA[20] which proceed in two steps. The first step consists in the estimations of ARIMA models correcting for deterministic effects (e.g. trading days or holidays). This step also implements an outlier detection procedure inspired by Chen and Liu (1993). The SA is then performed in the second step on the adjusted series.

For our purpose, we are only interested in the first step of the procedures although the second step confirms that no seasonal effects can be detected in our time

[18] For Venezuela, the UNHCR ODP provides information in cooperation with R4V (Inter-Agency Coordination Platform for Refugees and Migrants from Venezuela) https://www.r4v.info/en/regional. Breakdown by country of destination is not available on ODP but retrievable from R4V PDF updates on population in the region.

[19] https://ec.europa.eu/eurostat/cros/content/software-jdemetra_en

[20] These methods are the *X12-ARIMA* procedure used by the U.S. Census Bureau and the *TRAMO-SEATS* procedure developed by the Bank of Spain. The two procedures differ in the estimation of the seasonal component which is performed using Moving Averages (*X12-ARIMA*) or unobserved component models (*TRAMO-SEATS*).

series. We consider two types of outliers (the additive outlier and the level shift outlier, see Chen and Liu (1993)) and include conflict and disaster events up to two lags as additional exogenous variables in the ARIMA specification. We then retrieve potential outliers using JDemetra+, check the validity of the specification using numerous statistical tests offered by the software (e.g., normality and absence of autocorrelation in the residuals) and correct our original series accordingly. Furthermore, the software estimates missing values which provides an alternative to the linear interpolation used in the main specification. The procedure leads to the correction of few outliers and Figure 4 and Figure 5 in the Appendix display the adjustment for some selected series.

Results of the model are similar to those of the main specification, with lags of violent events due to conflict being all significant at 1% level, except for the fifth (see Table 8 in the Appendix). As for the main specification, the response of displacement to conflict is hump-shaped with the first lag reporting the highest coefficient. Lags of mass disasters, instead, are not found to be significant.

6 Conclusions

The number of forcibly displaced people almost doubled between 2010 and 2020, with cross-border displacement representing 40% of all displacement. With **neighbouring countries hosting about 75% of all people displaced across borders**, forced displacement disproportionately affects some regions more than others, and developed countries play only a minor role in worldwide support to people in need. Given the importance of neighbouring countries in hosting displaced populations, understanding the development of cross-border displacement due to conflict- or disaster- induced instability is crucial. This evidence can help **formulate adequate and timely policy responses** to support host countries and international organisations working on the ground: not only providing financial and technical support, but also proactively planning availability in refugees camps, humanitarian corridors and resettlement pledges. Finally, understanding the time sequence of cross-border displacement into neighbouring countries is important to assess **possible developments of mixed-migration movements along the routes leading to the EU**.

Using monthly observations for seven countries of origin and twenty-one neighbouring countries, this analysis provides evidence on the effects of violent events due to conflict and mass disasters at origin on forced displacement into neighbouring countries. The analysis applies a FDL model in first-differences to estimate which - and how many - time-lags of both conflict and disasters report a statistically significant coefficient explaining changes in displaced people. The AIC and BIC information criteria are used for the selection of a parsimonious lag order specification, which nonetheless fits the data well enough.

Lags of violent events due to conflict always report a significant coefficient while this is not the case for mass disasters. In terms of number of lags (i.e., months) for which these effects on displacement into neighbouring countries are significant, this appears to be **between the first and fifth lags after an increase in violent events due to conflict at origin**. Moreover, the coefficients of the different lags seem to show a **non-linear trend, with the first lag reporting the largest coefficient** and then decreasing in the following months, with the fifth lag also reporting a lower level of statistical significance. Significant results for mass disasters are found only for the first and fourth lags.

The findings therefore provide evidence on the **higher significance of conflict compared to disasters in explaining first-time displacement into neighbouring countries**, and that their effects are persistent and can last for several months after the occurrence of the events themselves. It is important to keep in mind, however, that mass disasters are considerably less frequent than violent events due to conflict, but are found to have a larger impact on displacement in neighbouring countries when coefficients are significant. Limited significant results for disasters should not be overlooked, in fact, as disruption due to these events can exacerbate instability in origin countries and fuel conflicts, which, in turn, stimulate first-time displacement into neighbouring countries.

Last but not least, it is also important to differentiate between mass disasters impacting the population in a relatively short window of time from changes in weather conditions, which can be sustained and indicate an overall shift in climate conditions. While a flood can suddenly force people to move because of massive disruption of the concerned area, prolonged periods of droughts might progressively worsen living conditions and thus push people to move. Moreover, depending on the areas affected by worsening of weather conditions, the impact on displacement might differ: droughts along the paths towards neighbouring countries might for instance hamper crossing. The findings of this report are therefore important to better inform the modelling under the EMT in terms of estimated impact on displacement, its possible drivers (e.g. conflict, disasters, weather conditions) and their effects over time.

References

Studies

Abel, G. J., Brottrager, M., Crespo Cuaresma, J., & Muttarak, R. (2017, April). Examining the link between climate, conflict and cross-border migration using Heckman selection with gravity model. In *XXVIII IUSSP International Population Conference*.

Abel, G. J., Brottrager, M., Cuaresma, J. C., & Muttarak, R. (2019). Climate, conflict and forced migration. *Global environmental change*, *54*, 239-249.

Burke, M., Hsiang, S. M., & Miguel, E. (2015). Climate and conflict. *Annu. Rev. Econ.*, *7*(1), 577-617.

Chen, C., & Liu, L. M. (1993). Joint estimation of model parameters and outlier effects in time series. *Journal of the American Statistical Association*, *88*(421), 284-297.

Drukker, D. M. (2003). Testing for serial correlation in linear panel-data models. *The stata journal*, *3*(2), 168-177.

Enders, W. (2008). *Applied econometric time series*. John Wiley & Sons.

Hsiang, S. M., & Burke, M. (2014). Climate, conflict, and social stability: what does the evidence say?. *Climatic change*, *123*(1), 39-55.

Hsiang, S. M., Burke, M., & Miguel, E. (2013). Quantifying the influence of climate on human conflict. *Science*, *341*(6151), 1235367.

IDMC. (2019). *African Report on Internal Displacement.* https://www.internal-displacement.org/sites/default/files/publications/documents/201912-Africa-report.pdf

Im, K. S., Pesaran, M. H., & Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of econometrics*, *115*(1), 53-74.

Jahani, A., Arabnejad, H., Suleimanova, D., Vuckovic, M., Mahmood, I., & Groen, D. (2021, June). Towards a Coupled Migration and Weather Simulation: South Sudan Conflict. In *International Conference on Computational Science* (pp. 502-515). Springer, Cham.

Maddala, G. S., & Wu, S. (1999). A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and statistics*, *61*(S1), 631-652.

Moore, W. H., & Shellman, S. M. (2004). Fear of persecution: Forced migration, 1952-1995. *Journal of Conflict Resolution*, *48*(5), 723-745.

Naudé, W. (2010). The determinants of migration from Sub-Saharan African countries. *Journal of African Economies*, *19*(3), 330-356.

Ruyssen, I., & Rayp, G. (2014). Determinants of intraregional migration in Sub-Saharan Africa 1980-2000. *Journal of Development Studies*, *50*(3), 426-443.

Sax, C., & Steiner, P. (2013). Temporal disaggregation of time series. *The R Jurnal, 5:2*, 80-87.

UNHCR. (2017). *Global Trends. Forced Displacement in 2016.* https://www.unhcr.org/5943e8a34.pdf

UNHCR. (2021). *Global Trends. Forced Displacement in 2020.* https://www.unhcr.org/60b638e37/unhcr-global-trends-2020

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Cengage learning.

Databases

UNHCR Operational Data Portal (ODP)

https://data2.unhcr.org/en/situations

UCDP Georeferenced Event Dataset (GED)

Pettersson, Therese, Shawn Davis, Amber Deniz, Garoun Engström, Nanar Hawach, Stina Högbladh, Margareta Sollenberg & Magnus Öberg (2021), Organized violence 1989-2020, with a special emphasis on Syria, *Journal of Peace Research*, 58(4).

Sundberg, Ralph and Erik Melander (2013), Introducing the UCDP Georeferenced Event Dataset, *Journal of Peace Research*, 50(4), see https://ucdp.uu.se/downloads/index.html#ged_global.

Emergency Events Database (EM-DAT)

EM-DAT, CRED / UCLouvain, Brussels, Belgium – (D. Guha-Sapir), see https://www.emdat.be/

Rulers, Elections, and Irregular Governance (REIGN) dataset

Bell, Curtis, Besaw, Clayton., Frank, Matthew. (2021). The Rulers, Elections, and Irregular Governance (REIGN) Dataset, Broomfield, CO: One Earth Future, see https://oefdatascience.github.io/REIGN.github.io/.

FAOSTAT

https://www.fao.org/faostat/en/#home

World Development Indicators (WDI)

https://datatopics.worldbank.org/world-development-indicators/

Appendix

Countries and time coverage

Table 4 Countries and time coverage

Country of origin	Neighbouring country	Start date
Burundi	Rwanda	Jun-15
	Tanzania	Mar-15
	Uganda	Feb-18
Central African Republic	Cameroon	Feb-13
	Chad	Dec-12
	Congo	Dec-13
Democratic Republic of the Congo	Burundi	Feb-18
	Malawi	Feb-18
	Rwanda	Jan-18
	South Sudan	Jan-18
	Tanzania	Jan-18
	Zambia	Jan-17
	Zimbabwe	Feb-18
Nigeria *	Total *	Apr-14
Somalia	Ethiopia	Jan-13
	Uganda	Feb-18
South Sudan	Democratic Republic of the Congo	Oct-14
	Ethiopia	Dec-13
	Kenya	Dec-13
Syria	Egypt	Jun-12
	Iraq	Jan-12
	Jordan	Jan-12
	Lebanon	Apr-12
	Turkey	Jan-12

* For Nigeria, the breakdown by neighbouring country is not available, so the total displacement in Niger, Chad and Cameroon is used.

Results

Figure 3 Information criteria for a maximum number of lags set to twelve

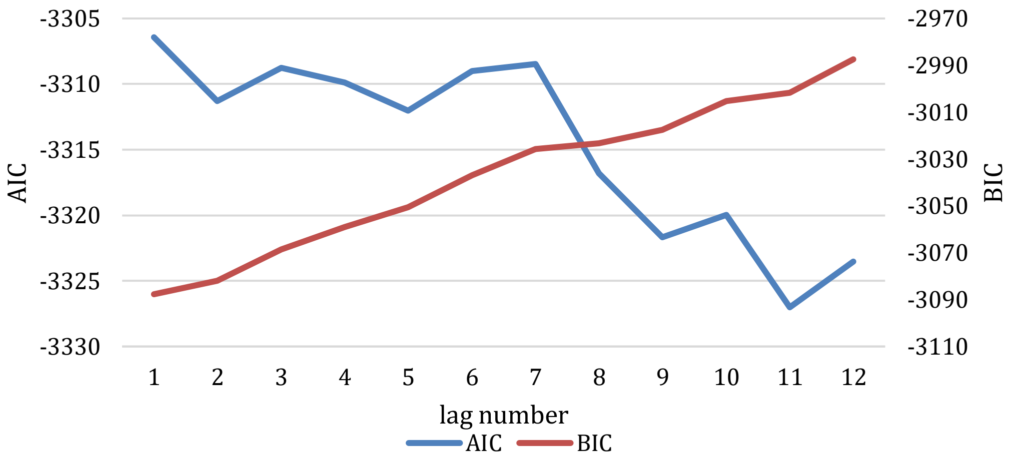


Table 5 Full set of results – lags two to six

Number of lags	2	3	4	5	6
Conflict	0.00013**	0.00013**	0.00016***	0.00015***	0.00016***
	(0.00005)	(0.00005)	(0.00005)	(0.00005)	(0.00005)
Lag1	0.00015***	0.00016***	0.00018***	0.00019***	0.00019***
	(0.00005)	(0.00005)	(0.00005)	(0.00005)	(0.00005)
Lag2	0.00012**	0.00013***	0.00014***	0.00016***	0.00017***
	(0.00005)	(0.00005)	(0.00005)	(0.00005)	(0.00005)
Lag3		0.00007***	0.00009***	0.00010***	0.00011***
		(0.00002)	(0.00002)	(0.00002)	(0.00002)
Lag4			0.00011***	0.00013***	0.00013***
			(0.00003)	(0.00003)	(0.00003)
Lag5				0.00008*	0.00009**
				(0.00004)	(0.00004)
Lag6					0.00003
					(0.00004)
Climate	-0.00283	-0.00280	-0.00192	-0.00239	-0.00227
	(0.00394)	(0.00395)	(0.00386)	(0.00378)	(0.00379)
Lag1	0.00493	0.00528	0.00687*	0.00654*	0.00648*
	(0.00320)	(0.00328)	(0.00365)	(0.00367)	(0.00372)
Lag2	-0.00135	-0.00118	0.00095	0.00023	0.00027
	(0.00396)	(0.00389)	(0.00409)	(0.00420)	(0.00420)
Lag3		0.00052	0.00325	0.00279	0.00278
		(0.00206)	(0.00241)	(0.00250)	(0.00272)
Lag4			0.00598**	0.00493*	0.00503*
			(0.00254)	(0.00281)	(0.00292)
Lag5				-0.00218	-0.00219
				(0.00231)	(0.00234)
Lag6					-0.00002
					(0.00321)
Population (log)	5.26982*	4.91720*	4.06691	3.76195	3.67323
	(2.77018)	(2.76247)	(2.80491)	(2.86255)	(2.84119)
Unempl. rate	0.04711	0.04302	0.02889	0.02339	0.02101
	(0.18068)	(0.18038)	(0.18036)	(0.18033)	(0.18020)

GDPpc (log)	6.61020*	7.22956*	8.46906**	9.90196**	10.34877**
	(3.90359)	(3.88533)	(3.92498)	(4.17063)	(4.10325)
$GDPpc^2$ (log)	-0.50036	-0.58207	-0.74715*	-0.92603**	-0.98289**
	(0.42254)	(0.41967)	(0.42528)	(0.45650)	(0.44856)
Constant	0.18636***	0.17338***	0.14253***	0.12065***	0.11460***
	(0.03831)	(0.03672)	(0.03654)	(0.04033)	(0.03944)
Observations	1,688	1,688	1,688	1,688	1,688
R-squared	0.197	0.199	0.207	0.211	0.212
Adjusted R-squared	0.176	0.178	0.185	0.187	0.187

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6 Conflict and Disasters – five lags specification including Venezuela

Lag number	*Conflict*	*Disasters*
0	0.00015***	-0.00304
	(0.00005)	(0.00364)
1	0.00020***	0.00620*
	(0.00005)	(0.00359)
2	0.00016***	0.00040
	(0.00005)	(0.00406)
3	0.00010***	0.00224
	(0.00002)	(0.00245)
4	0.00013***	0.00477*
	(0.00003)	(0.00268)
5	0.00008*	-0.00207
	(0.00004)	(0.00220)
Observations	1,800	
R-squared	0.205	
Adjusted R-squared	0.180	

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7 Conflict and Disasters – five lags specification – AR(1)

Lag number	*AR(1)*	*Conflict*	*Disasters*
0		0.00015***	-0.00235
		(0.00005)	(0.00378)
1	0.15503***	0.00017***	0.00674*
	(0.04722)	(0.00005)	(0.00370)
2		0.00013***	-0.00060
		(0.00005)	(0.00421)
3		0.00008***	0.00268
		(0.00002)	(0.00244)
4		0.00012***	0.00478*
		(0.00003)	(0.00278)

5		0.00006 (0.00004)	-0.00284 (0.00233)
Observations		1,688	
R-squared		0.233	
Adjusted R-squared		0.209	

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 4 Adjusted series of displaced population with JDemetra+ software- Central African Republic to Chad

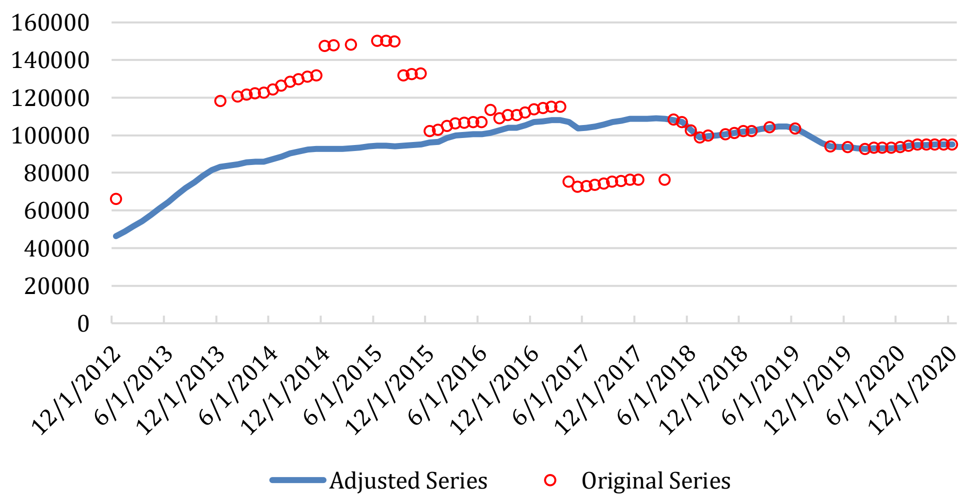


Figure 5 Adjusted series of displaced population with JDemetra+ software- South Sudan to Kenya

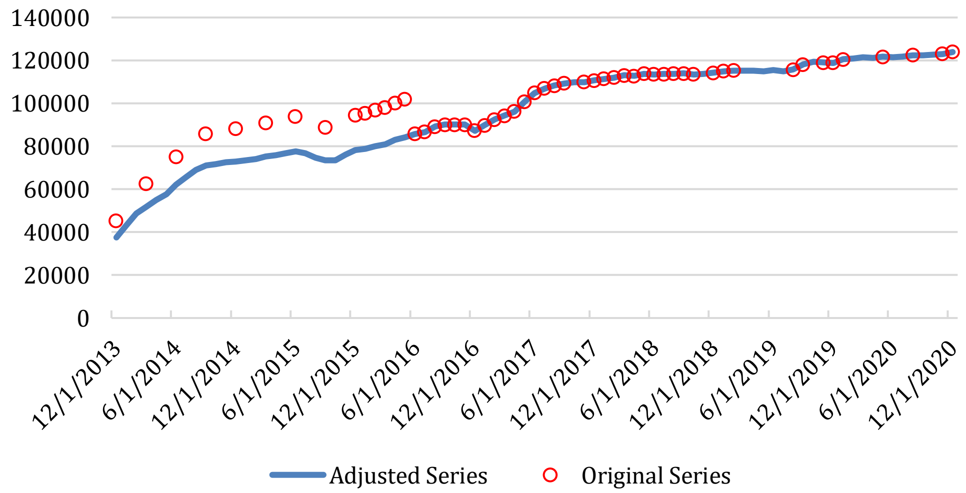


Table 8 Conflict and Disasters – five lags specification – Seasonal and outlier adjustment (JDemetra+)

Lag number	*Conflict*	*Disasters*
0	0.00006* (0.00003)	-0.00069 (0.00141)

1	0.00012***	-0.00099
	(0.00003)	(0.00152)
2	0.00009***	-0.00083
	(0.00003)	(0.00165)
3	0.00007***	-0.00111
	(0.00003)	(0.00186)
4	0.00007***	-0.00078
	(0.00002)	(0.00156)
5	0.00003	-0.00034
	(0.00003)	(0.00117)
Observations	1,690	
R-squared	0.395	
Adjusted R-squared	0.377	