Hertie School

High risk migration via the Central Mediterranean: understanding causes of flight and gender heterogeneity

by

Ana Elisa von Borstel da Silva

Thesis submitted in partial fulfilment of the requirements for the degree of Master of Public Policy

Supervised by Prof. Dr. Julian Wucherpfennig

Berlin, May 2023

Table of Contents

Acknowledgements

I would like to express my sincere gratitude to my thesis supervisor, Prof. Dr. Julian Wucherpfennig, for the insightful conversations, guidance, patience, and constructive feedback. Prof. Wucherpfennig inspired me to research a real-world problem that I am passionate about, despite all the challenges related to the lack of available, complete, and reliable data. It was a great privilege and honor to study and work under his guidance.

I would like to extend my gratitude to the IPLI Foundation and to the Centre for International Security for the grant that supported my research.

I am extremely grateful to my master's colleagues, Katarina Vujovic, Vartika Savarna and Nino Kvaratskheliia, for all the late nights at the library, the Saturdays we spent solving Statistic II assignments, and the uncountable amazing adventures in Berlin. You are such intelligent, successful, and inspiring women. My life in Berlin would not have been the same without you.

Nobody has been more important to me in the pursuit of this degree than the members of my family. I would like to specially thank my parents Rosi and Humberto and my sister Cristiane for always believing in me and supporting my dreams. I am so lucky to have such a loving and caring family! You are my rock and strengthen. I would like to also thank my sister's husband Fernando, my nephews Henrique and Heitor, my grandmothers Anita and Geni, my aunties Heloisa and Helenice, and my godfather and godmother Alberto and Ivone. *Saudade enorme de vocês*.

Finally, I would also like to thank my childhood friends Elisa, Giovanna, Camila, Rodrigo, and Cecilia. Even so far away, you were always there to listen (or better, read) the struggles of living in Germany and completing my master's degree. Thank you for the videocalls, for the supportive messages and for always being there for me.

Executive Summary

This research aims to investigate the main causes of irregular migration to Europe via Central Mediterranean in the last years. Historically, the Mediterranean Sea has always been used as a migratory route between South Europe, North Africa, and Middle East. There are records of missing and/or dead people trying to reach Europe via Mediterranean since 1993[1]. However, from 2014 this route has become increasingly used to reach Europe irregularly, which resulted not only in an elevated number of dead and missing people, but also in a political turmoil in Europe. Understanding how events happening in other countries can affect the flow of migrants using the Central Mediterranean would allow Europe to design better policies to address the irregular crossings and the increasing number of fatalities in the Mediterranean.

By using arrival data in Italy, geolocated conflict data, and economic and political indicators, this research provides evidence on the push factors of irregular migration via the Central Mediterranean. Fixed effects regressions are used to analyze how changes in the context of countries in Africa, Southern Asia, and Western Asia, affected the number of migrants and asylum seekers arriving in Italy from 2011 to 2018. Evidence is found to support that violence is one of the main drivers of migration flow via Central Mediterranean.

The second part of this research focus on how causes of flee affect women and men differently in their migration behavior. Most existing theories and empirical research on aggregated level of causes of forced migration do not account for gender heterogeneity (Schmeidl, 1997; Davenport et all, 2003; More & Shellman, 2007; Melander & Öberg, 2007). However, there is evidence, mainly from qualitative studies, that women and men are differently affected by violence and conflict, as well by other economic, political, and social factors in their decision to migrate (Aksoy & Poutvaara, 2019; Blangiardo, 2012; Piper, 2006; Müller-Funk, 2019). For example, men are more at risk to be forced to join the military service or armed groups. Women, on the other hand, might not have the financial or social means to flee the country of conflict to Europe or they might have to migrate with their children, which makes long journeys more difficult.

The results show that the presence of one-sided conflict and the number of civilian fatalities are good predictors of arrivals in Italy. Regarding economic factors, there is not enough evidence to confirm the causal relationship between poverty/unemployment and the migration flow. Political indicators, like political rights and women empowerment index, also seem to have an effect on arrivals in Italy. Evidence is found to support that migration timing and flows of women and men are differently affected by conflict.

This analysis contributes to the literature review on aggregated macro level causes of forced displacement by analyzing a recent event in Europe (the 2015 Refugee Crisis), by using a new data source (arrivals in Italy) and by bringing a gender perspective to it. My conclusion and policy recommendation are that violence is an important factor of the European migration crisis and the peace component should be incorporated to international development projects addressing root causes of migration. Moreover, a gender perspective should be incorporated into those projects since evidence shows that police can affect gendered patterns of migration.

[1] See UNITED for Intercultural Action – campaign 'Fortress Europe No More Deaths' unitedagainstrefugeedeaths.eu – listofdeaths@unitedagainstracism.org

Introduction

From 2014 Europe experienced one of its biggest refugee/migration crises since the breakup of Yugoslavia in the 1990's. In 2015 more than one million refugees, asylum seekers and migrants arrived in Europe by sea and land, crossing the Balkans or the Mediterranean from Turkey to Greece and from North Africa to Italy and Malta[2]. Although the number of migrants crossing the Mediterranean decreased in the following years, Europe still receive more than 100,000 arrivals per year[3]. The number of arrivals in 2022 were higher than 2018 and 2019, showing that this crisis is far from being solved.

The crisis Europe face is not only a migration one. The situation has strong humanitarian and political components. The Central Mediterranean route is the deadliest irregular route in the world, with the number of dead and missing people counting at least 17,000 casualties since 2014[4]. It is likely that this number is much higher, since it is not possible to collect data on all crossings and incidents in the Mediterranean, due to the illegal nature of the crossings.

Besides the incidents in the Mediterranean, there are records of human rights violations originating from European Union agreements with Turkey, Libya, and Morocco. Arbitrary detention of people attempting to migrate to Europe, pushbacks from costal guards, and shooting of migrants in the border regions are some examples of human rights violations (Niemann & Zaun, 2018).

Regarding the political component, right wing and conservative politicians gained space in national and European politics by advocating for measures to increase border control and restrict migration to Europe. Southern European countries were highly affected by the crisis for being the countries where asylum seekers entered Europe. The lack of a coordinated EU response and the weaknesses of the Dublin Regulation led to border closures inside the Schengen area, humanitarian crisis in the Greek Islands and Italy coast, and a threat of European cohesion and stability. The refugee crisis was one of the main components that culminated in the United Kingdom leaving the European Union, and politicians of other countries advocating for the same.

Despite the complexity and dimension of the 2015 refugee crisis, this research focuses on how events and changes in conditions in Africa, Southern Asia, and Western Asia affected the arrivals via the Central Mediterranean route. There are methodological and data constrain reasons for restricting this analysis to the Central route. First, the data used in the gender component of this analysis is only available for arrivals in Italy. Second, the percentage of migrant men crossing the Mediterranean in comparison to migrant women is much higher for the central route than to the eastern one.

The Central Mediterranean route is mainly used by migrant men, with women accounting for only 20% of the migrants. In some periods of time, there were more accompanied and

[2] See UNHCR Operational Data Portal, Mediterranean Situation. Available at https://data.unhcr.org/en/situations/mediterranean. Access on 24th April 2023
[3] See UNHCR Operational Data Portal, Mediterranean Situation. Available at https://data.unhcr.org/en/situations/mediterranean. Access on 24th April 2023
[4] See UNHCR Missing Migrants Project. Available at https://missingmigrants.iom.int/region/mediterranean

unaccompanied minors using this route than women. The same is not observed for the Eastern Mediterranean route, where a higher percentage of women and children use the route[5].

It is not clear though what are the causes of this difference in shares of men and women using the Central and Eastern routes. It could be that women are more restricted in their migration decision-making because of moral, socio and economic constrains in comparison to men. It could also be that the Central Mediterranean route is too dangerous, and women are less risk averse, or they are at more risk of suffering gender violence in their journey. Alternatively, women might respond differently to violence, political and economic factors than men, with women more willing to use a dangerous route to flee violence and conflict in comparison to migrating due to economic and political reasons. Unfortunately, gender disaggregated arrival data in Greece is not available, which does not allow for migration pattern comparisons between the central and eastern routes. However, since the Central Mediterranean route is a mixed flow of asylum seekers and economic migrants, it is possible to present some evidence on how conflict and other factors affect women and men differently.

Another reason for not including arrival data in Greece in this analysis is because there is a higher variation of migrants from different nationalities using the central route, in comparison to the dominance of Syrians and Afghanis using the eastern route. Since the main nationality using the Eastern Mediterranean route was Syrian, the results of the effect of changes in conditions in other countries of migrants using this route would be downsized because of the magnitude of the Syrian conflict. Therefore, using the central route allows for more variability of nationalities of migrants crossing the Mediterranean. Moreover, after April 2016, the number of arrivals in Greece decreased to less than 3,000 arrivals per month, whereas in Italy the monthly number of arrivals continue stable at more than 10,000 arrivals until August 2017. Nonetheless, the severity of the Syrian conflict is also captured by the arrival data in Italy.

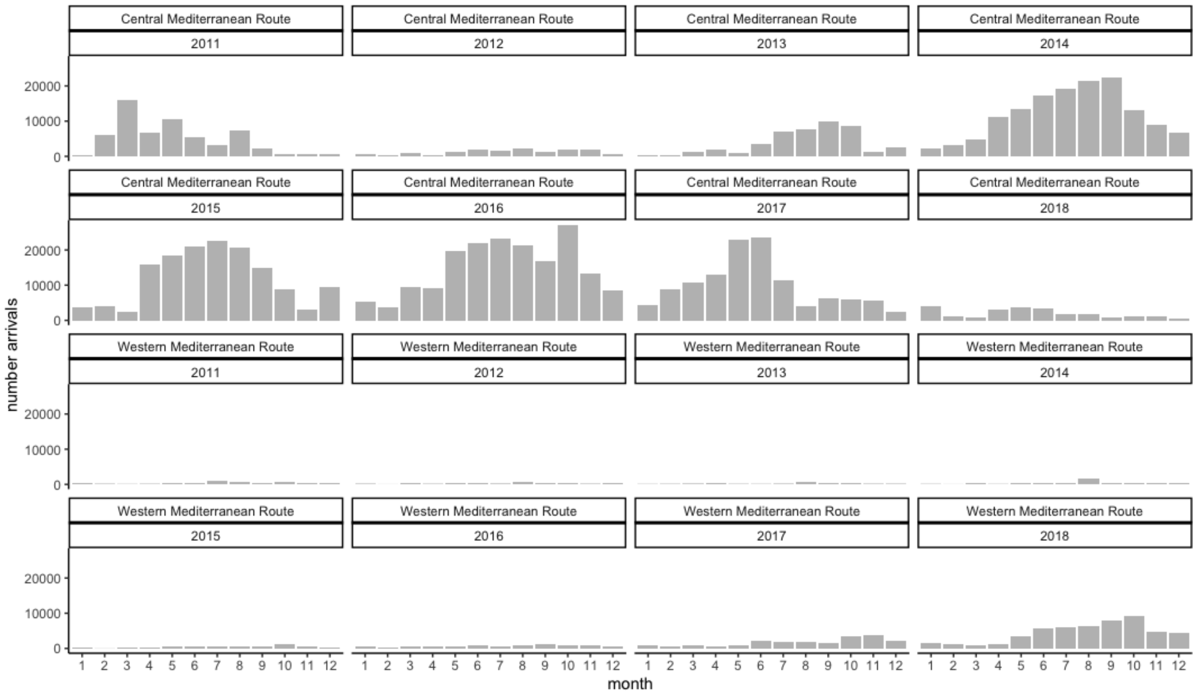
Finally, with the start of the Syrian conflict in 2011, the Central Mediterranean route became the main route of irregular migration from Africa to Europe. This is due to the increase difficulty of crossing the Syrian territory to arrive in Turkey, where the Eastern Mediterranean and the Balkan routes are alternatives to migrate to Europe. Therefore, the Syrian conflict created a natural experiment in which the Central Mediterranean route became the main irregular route to access Europe from Africa, which allows to observe how violence, economic conditions and political factors in the continent affected the irregular migration patterns. Moreover, the increase of arrivals of Syrian refugees in Greece, the closure of borders of countries from the Balkan route, the EU-Turkey agreement, and the difficulties of moving from the Greek islands to the continent made the Central Mediterranean route an important alternative for migrants from Southern and Western Asia.

Figure 1 shows the monthly number of arrivals of the Central and Eastern Mediterranean routes from 2011 to 2018. The Western Mediterranean route (calculated by arrivals in Spain) accounted for only a small number of arrivals in Europe until 2018. It is not clear why the Central Mediterranean route is preferred in relation to the Western route. Nonetheless, this research focus on the Central route only since it is the main route used, and it is the only route where arrivals disaggregated by gender and nationality are available. This data is important for the second part of this research, in which the gender component is analyzed. Therefore, for

[5] See UNHCR Operational Data Portal, Mediterranean Situation. Available at https://data.unhcr.org/en/situations/mediterranean. Access on 24th April 2023

better comparison between models, arrivals in Spain or in Greece are not included and the period of analysis is restricted from 2011 to April 2018.

Figure 1: Monthly arrivals via Central and Eastern Mediterranean Routes from 2011 to 2018



The focus of this paper, therefore, is on push factors that influenced migration flows via the Central Mediterranean toward Europe and how these factors affected men and women differently in their migration behavior and timing. The motivation questions are: what are the main causes of irregular migration via Central Mediterranean route? What are the responsible changing conditions of countries producing migrants using this route? How push factors affect men and women differently in their migration behavior and timing?

These questions are important to understand how conflict, economic and political factors in Africa, Southern Asia, and Western Asia are related to the Mediterranean migration crisis and what are the best approaches for addressing this crisis. Europe's responses to the crisis involved internal and external measures, such as border control, fighting smuggling of migrants and human trafficking and investing in development projects in Africa (Niemann & Zaun, 2018). One example is the EU Emergency Trust Fund for Africa, which focus on creating economic and employment opportunities in selected countries and improving migration management[6]. This initiative is part of European Union and international development agencies strategy to finance development projects in Africa and Asia to address the root causes of migration.

The aim of this research is to bring evidence to the political debate of international development projects working as a mechanism to address the root causes of forced displacement. Some authors highlight that there is no causal relationship between international development and international migration (Brück, 2016; Schmeidl, 1997) and that the development and migration nexus are rather a political construction (Greiger & Pécoud, 2013). There is mixed evidence that development projects reduce or even increase international migration. Although this topic

[6] See European Union Trust Fund for Africa. Available at <https://trust-fund-for-africa.europa.eu/index_en> Access on 24th April 2023

has increased in importance in the last decades, not many empirical studies has been published so far and project evaluations are marked by mainly anecdotal stories or monitoring data.

This research also contributes to the debate on how migration behavior is affected by gender. There are few articles using empirical research to analyze the role gender plays in migration patterns. Most studies focus have a qualitative approach or use a case study to draw conclusions on how women are affected by migration. The classical literature on macro-level analysis on causes of forced displacement do not account for gender differences in migration flows, as it is discussed in the following section.

This paper is structed as follows: in the first part research and prior studies on the topic of causes of forced displacement and gendering migration are discussed. Then, the research design, variables, method, and hypothesis are presented. In the third part, descriptive results are explored, following by the analysis of the regression results. Finally, the conclusion is presented, and policy recommendations are drawn.

Literature review

Prior studies on forced displacement offer evidence that conflict is the one of the main causes of flee (Schmeidl, 1997; Adhikari, 2013; Bohra-Mishra & Massey, 2011; Brück et all, 2018). When a conflict starts in a country, involving either State or non-State actors, citizens might be affected not only by the deterioration of economic opportunities or in their movement of freedom, but also might feel that their personal security and integrity are at risk. There are several macro country level studies linking different types of conflict (civil war, genocides, politicide, dissident conflict) with internally displacement and refugee flows (Schmeidl, 1997; Davenport et all, 2010; Moore & Shellman, 2004). Micro-level studies show similar results for how violence and fear of violence affect the decision to flee (Adhikari, 2013; Bohra-Mishra & Massey, 2011).

Other causes of forced displacement, such as economic factors and human rights violations for example, do not share the same consensus from the literature review. There is mixed evidence on how economic deterioration, type of political institution, and political persecution affect forced displacement flows and intensity. Literature on voluntary migration usually focus on the influence of push and pull factors on migration patterns. In this ramification of research, evidence is found for economic factors working both as push and pull factors for migration dynamics.

Since this research focus on migrants crossing the Mediterranean, which it is a mix flow of migrants from different countries and with different motivations to migrate, literature on forced and voluntary migration are explored. Notwithstanding, the focus of this analysis is not on pull factors of migration, but rather on how changes in conflict, economic situation and political context in Africa, Middle East and South Asia affected the flow of migrants crossing the Central Mediterranean.

A classical article about causes of forced displacement using aggregated country longitudinal data is from Schmeidl (1997). The author identifies three main factors in the dynamic of forced displacement: root causes (mainly economic factors), proximate conditions (mainly political factors), and intervening factors (facilitators or obstacles). The pooled time-series analysis is conducted for countries from the global South from 1971 to 1990. The results show a statistically significant effect of genocide, politicide, and civil war on refugee stock change (what the author uses as a proxy for migration flow). On the other hand, the author did not find a direct effect of poverty and underdevelopment on refugee migration. The author concludes that political violence is a better predictor of refugee migration than economic factors.

Similar results are presented by other authors. Moore and Shellman (2004) argue that people closely monitor the violence situation in their country and use this information in their decision to flee or stay. Publicly available information influences aggregate outcomes, meaning that news on violence affect the aggregate outcome of people leaving the country. The authors find that violence produced either by government or dissidents are the primary causes of forced displacement, whereas institutional democracy and GNP per capita have only a small predictor power. Adhikari (2013) finds similar results for individual level data, by addressing how not only actual violence is a cause of displacement, but also the perceived threat of violence. By using a choice-centered approach to study forced migration in Nepal, the author finds that violence and the threat of violence have a statistically significant effect on the individual’s decision to flee.

Melander and Öberg (2007) study how variation in magnitude and scope of fighting across armed conflicts affect the intensity of migration flow. The authors find no correlation between the intensity of the conflict, measured by number of battle-related deaths, and the flow of forced migrants. However, the authors find an effect of the geographical scope and location of the conflict, meaning that conflicts spread to bigger and more urban areas have a higher effect on the flow of forced migrants.

Adhikari (2013) arguments that survival and deterioration of economic conditions are also important determinants of forced migration, since in economically precarious societies, individuals’ economic survival might be compromised by economic breakdowns caused by conflicts. Therefore, not only violence and perceived violence are important factors in the decision to flee, but as well the economic conditions of a place.

Indeed, evidence related to the effect of economic factors on forced and voluntary migration are mixed, with some authors finding a positive effect, others finding a negative one, and some finding no significant effect. Moore and Shellman (2004) find a negative impact of GNP per capita on the number of forced migrants produced by the countries of analysis, meaning that higher GNP per capita reduces the number of refugees produced by a country. Weber (2019) finds no significant effect of GDP growth on the arrivals in Europe.

Kang (2021) conduct a standard gravity model to analyze push and pull factors responsible for the European refugee migration from 2008 to 2014. The results show that GDP per capita of countries of destination and origin are statistically significant predictors of migration, meaning that GDP per capita of countries of origin have a negative effect on the number of asylum applications in Europe, whereas countries of destination have a positively effect. However, coefficients for unemployment rates of countries of origin are not statistically significant, whereas unemployment in countries of destination is positively associated with asylum applications. The author concludes that it is not possible to affirm that refugee flow is driven by economic pull factors, such as seeking job opportunities in Europe. Finally, the author finds that political stability is the most significant variable in explaining the outflow of asylum seekers from countries of origin.

Brück et all (2018) also analyze pull and push factors related to the Arab Spring migration towards Europe from 2006 to 2016. Little evidence is found for economic variables (GPD per capita) playing a pull or push factor in the number of asylum applications. The authors find, however, a positive impact of employment rate, but concludes that violence is a major driver of the decision to flee.

Neumayer (2005) analyses the root causes of forced migration to Europe from 1982 to 1999. The author finds evidence that economic factors are related to the number of asylum application in Europe, with economic growth rate being positively associated with the number of asylum applications. However, it is found that the relationship between GDP per capita and migration has an U shape than a linear format. The author also finds substantive effect of human rights violation, dissident violence, civil war and state failure on migration to Europe.

Moore & Shellman (2007), on the other hand, find evidence that refugees are no bogus migrants, as defended by Neumayer (2005). The authors find significant differences between the flow of refugees between bordering countries and countries farther away. Border war and bordering civil wars have a stronger effect on refugee flows than average wages and colonial ties.

Studies that use political and other control variables in their analysis also find mixed results of significancy and effect size. Kirwin and Anderson (2018) in their study on reasons for West African migration towards Europe find that migrants prepare their journey carefully and with years in advance. It is found that economic standing alone does not have a significant effect on Nigerians' desire to migrate, but rather a lack of sufficient security or trust in the political system. The strongest relationships found was satisfaction with democracy and trust in the police.

Network effects and temporal dependency are also explored by other authors. Using country-dyads analysis of asylum flows from 48 African countries to 31 European countries, Weber (2019) finds that, although an increase in intensity and quantity of violent conflicts had an effect on the number of asylum applications in Europe, violent conflicts alone do not explain the sudden high number of arrivals in Europe in 2015. The author further explores migration as a self-reinforcement process for the period after 2015.

Brück et all (2018) find a positive effect of network (measured as the lag of asylum applications) on the number of asylum applications in countries of destination. The mechanism behind is that migrants have more access to information about the destination countries when other migrants from the same nationality already live there. The network effect is similar for forced and voluntary migration. Weber (2019) uses the lag of asylum seeker applications to control for network. Davenport (2003) and Schmeild (1997) use the lag of the dependent variable to account for related events over time.

Although literature and prior studies of the causes of forced migration offer evidence on how violence is one of the main causes of displacement and on how economic factors might play a role in the intensity of the migration flow, little evidence is presented on how these variables affect men and women in different ways and intensities. For example, Schon (2019), in his research on violence and migration timing, finds that women and men have different migration timing patterns, but does not further explore the causes of this difference.

Literature on gender and migration is heavily focused on qualitative and descriptive studies. Pier (2006) argues that is important to study the gendering of migration because men and women can be differently affected by migration laws and policies, which results in gendered patterns of migration. The author presents three main mechanisms that explains this difference: gender-segregated labor markets, gendered socioeconomic power structures and sociocultural definitions of appropriate roles in destination and origin countries.

For the Central Mediterranean, it seems plausible that the socioeconomic and sociocultural structures are partially shaping the gendering migration of the route. Aksoy and Poutvaara (2019) find that the shares of female migrants using the Central and Eastern Mediterranean routes in 2015 and 2016 to reach Europe are lower from countries in which traditional gender roles are stronger. However, the authors argue that women were more likely to migrate from countries experiencing intense conflict or persecution, since the risk of staying in the country is higher than the risk of migrating.

The risk of migration is also explored by Gerard and Pickering (2013) who studies experiences of transit of Somali refugee women who traveled from North Africa to Malta. The authors explore how EU securitization policies exacerbate the violence refugee women experience while in transit. The authors bring two examples on how the risks of irregular border crossing have gendered dimensions. For example, not only women are more at risk of suffering sexual violence, but the provision of sexual services in the border regions as a way to negotiate the crossings have consequences for women’s health and unwanted pregnancy. Moreover, the authors found evidence that women, in comparison to men, are more likely to be placed at more vulnerable seats on the boats crossing the Mediterranean.

Massey et all (1993) states that the New Economics of Migration theory understand the migration decisions are not individual, but rather shaped by large units of related persons, such as families or households. In developing countries, international migration to developed countries with higher wages can be seen as a strategy to mitigate risks of poor households.

The predominantly composition of migrant men using the Central Mediterranean route could also be explained by the New Economics of Migration. Since the route is dangerous and expensive, it could be that families understand that sending a young man to do the journey is the best strategy. Not only these young men could financially support their families once they migrate to Europe, but they could also facilitate other members of their families to migrate to Europe by family reunion and through safer pathways.

Blangiardo (2012) highlights that while economic motives largely explain the migration of males in the Mediterranean and Sub-Saharan countries, the reasons for female migration are more diverse, with marriage and family reunification playing an important role. Nonetheless, the author highlights a feminization of migration in the recent decades. In Mali, for example, although the proportion of men and women internally migrating is similar, men still dominate international migration patterns (with the proportion of male migrants twice the number of female migrants). Tunisia presents similar patterns: concerning international migration, men represent 75% of total stock of migrants and the main reasons for emigrating were economic factors and employment, whereas the main reasons of female migration were family unification and marriage. Similar patterns were found for Egypt, Jordan, Palestine. Some countries of the analysis, however, presented other factors for female migration, such as employment opportunity, study and autonomy/emancipation. In Senegal, the main reasons for female migration were indicated as work, family, marriage, and study. In Morocco, economic opportunities in destination countries were indicated as one of the main reasons of migration.

Nonetheless, Müller-Funk (2019) highlights the agency women have in shaping the family decision to migration. Using a mixed-method study of surveys and interviews with Syrian refugees living in Turkey, the author finds that life satisfaction, hope for peace in Syria, access to work and sense of common cultural belonging are important factors shaping micro-level decision to migrate. The author finds that although the decision to migrate is usually a whole-family decision, women often exercise agency by influencing the destination and migration timing.

Theoretical Framework

The data for this analysis comes from different sources. The dependent variable is number of monthly arrivals in Italy, disaggregated by country of origin and gender (when available). The data for arrivals in Italy via the Central Mediterranean route was collected by Italian authorities, and compiled and published by Frontex, IOM and UNHCR. The data for total number of arrivals by country of origin comes from Frontex. The gender disaggregation data comes from IOM Displacement Tracking Matrix (DTM) for the years of 2015, 2016, 2017 and from UNHCR for the year of 2014. Gender disaggregation is not available for the years before 2014 and it is not available for all countries producing migrants crossing the Mediterranean. Arrival data from IOM dataset is in cumulative format, therefore data transformation is required to calculate monthly arrivals. If the cumulative arrival of a month is coded as zero, the monthly arrival for the previous month is also marked as zero. If the total number of arrivals for a month from the Frontex dataset is zero, the gender disaggregation arrival was also coded as zero. Further details on data cleaning and manipulation are presented in the appendix.

Arrival data is a good proxy for migration flow, for it measure how many people arrived in Italy in the period of analysis, despite their migration status. It includes not only asylum seekers and refugees, but also irregular and undocumented migrants, which is not captured by other sources of migration data. Most research on aggregated causes of migration use refugee stock or asylum applications to measure migration flow; however, these variables fail to account for irregular migrants. Other limitation of using stock to estimate flow migration is that stock data can be affected by naturalization and mortality rates. Therefore, arrival data provides a clearer picture of how many migrants crossed the Mediterranean when, which makes it more sensitive to events happening in the countries of analysis.

There are some limitations nonetheless: the number of fatalities related to migrants attempting to cross the Mediterranean is estimated to be high, therefore number of arrivals in Italy is very likely smaller in comparison to the number of migrants trying to reach Europe. However, it is likely that the risk of dying is the same for all migrants, regardless of their country of origin or gender, which would not cause selection bias in the model. Therefore, the number of arrivals in Italy still seems to be the best measure of flow migration to Europe.

In this research, forced and voluntary migrants are not distinguished, since the Mediterranean route has a mixed flow of refugees, asylum seekers and economic migrants. Without individual data, it is not possible to distinguish migrants based on their reasons to migrate. The data used for this analysis, aggregated arrival data in Italy, does not provide information on the reason to migrate of the migrants arriving in Italy, which makes the distinction impossible.

Notwithstanding, the underlying assumption of this research is that migrants would not use the deadliest irregular routes of the world unless the socio-economic and political conditions at their country of origin are critical. Migrating is a costly decision: leaving one's home, job, culture, social network, and family comes with a cost and the dangerousness of the route increases this cost. Therefore, this research treats the migrants crossing the Mediterranean mainly as non-voluntary migrants and uses the literature review on forced migration as a base, even though it acknowledges that pull factors are important to explain migration decision.

Countries from Africa, Southern Asia, and Western Asia not producing migrants crossing the Central Mediterranean are also included in the dataset to avoid selection bias. The final data includes 78 countries and 6,120 observations. Western Sahara, Seychelles, and Palestine were dropped from the sample because of missing observations of independent variables. Also, there are some limitations concerning the dependent variable and these countries. Since it is contested if Western Sahara and Palestine are considered states, it could be that migrants from these countries were registered as "Stateless" in the Frontex, UNHCR and IOM datasets. This is further discussed in the limitation of this research.

The independent variables are grouped in 3 main dimensions of causes of flee (violence, economic conditions, and political factors) and in obstacles/facilitators. For the variables related to violence and conflict intensity, data on individual events of organized violence from the UCDP Georeferenced Event Dataset Global version 22.1 is used. This dataset is UCDP's most disaggregated one, covering individual events of organized violence by month-year. Global development data comes from the World Bank Open Data and is used for the economic variables. The Quality of Government dataset, from the Gothenburg University, is used for the political variables.

The first dimension of causes of flee (violence) is measured by the absence or presence (and type) of conflict and by number of civilian casualties. There are 3 categories for type of violence at the UCDP dataset: state-based conflict, non-state conflict and one-sided violence. State-based conflict refers to conflicts between two parties of which one is a state actor; and non-state conflict to conflicts between two parties of which none is a state actor. One-sided violence refers to deliberated use of forced by state or non-state actors against civilians. The three types of violence are included in the regression model, since both state and non-state violence are related to human rights violation and forced migration (Moore & Shellmann, 2004; Davenport et all, 2003). Conflict data is available on a month-year unit of analysis.

The number of civilian fatalities is included in the model to account for conflict intensity and arbitrary violence spread. Schon (2019) explores how indiscriminate use of violence affects the decision to migrate. The author argues that people with advantaged social position might be able to protect themselves from target violence, but not from indiscriminate violence in residential areas. Melander and Öberg (2007) use total battle-related deaths (civilian and combatants) in their model but find no statistically significant result. Therefore, using civilian casualties instead of combatant deaths or total battle-related deaths aims to capture indiscriminate use of violence (especially in urban areas) instead of battle-related deaths that could occur in restricted/rural areas. Indiscriminate violence in urban areas is likely to affect at a higher level the aggregated number of people deciding to flee the country than battle-related violence. Melander and Öberg (2007) find that greater geographical scope of the conflict and urban areas have a significant effect on the flow of forced migrants.

The main hypothesis of this research is that violence is a good predictor of arrivals in Italy:

H1: the presence of conflict increases the number of migrants crossing the Central Mediterranean to Italy.

H2: the higher the number of civilian casualties*, the higher the number of migrants crossing the Mediterranean.*

The second dimension of causes of forced displacement (economic conditions) is measured by the GDP per capita and unemployment rate (in percentage of total labor force) of the countries of analysis. Data for GDP per capita comes from the World Bank and is used in the log format. Data for unemployment comes from the International Labour Organization Modelled Estimates and Projections database (ILOEST), also available at the World Bank Open Data. Both variables are displayed on a year unit of analysis. In some prior studies, variables to account for poverty are included in the analyses, such as energy consumption in kilograms of oil equivalent per capita (Schmeidl, 1997). However, I decided not to include energy consumption in my model because this variable is highly correlated to GDP per capita, which creates a problem of multicollinearity. Poverty rate is also not included in the model because data from the World Bank contain many missing observations for the countries of analysis, which reduces the sample to almost 50% of the original size.

Based on the literature review, I do not expect changes in economic conditions to be a good predictor of arrivals in Italy.

H3: changes in economic conditions of the country of origin does not have an effect on arrivals in Italy.

The third dimension of causes of flee (political factors) is measured by two variables available at the Quality of Government Dataset: political rights rate (from the V-dem Institute) and political terror scale (from the US State Department). The variables are presented on country-year unit of analysis and are described as the following:

"Political rights enable people to participate freely in the political process, including the right to vote freely for distinct alternatives in legitimate elections, compete for public office, join political parties and organizations, and elect representatives who have a decisive impact on public policies and are accountable to the electorate. The specific list of rights considered varies over the years. Countries are graded between 1 (most free) and 7 (least free)." (Standard Codebook, 2023, pp. 776)

"Political Terror Scale Levels from the the U.S. State Department Country Reports on Human Rights Practices:
1. Countries under a secure rule of law, people are not imprisoned for their view, and torture is rare or exceptional. Political murders are extremely rare.
2. There is a limited amount of imprisonment for nonviolent political activity. However, few persons are affected, torture and beatings are exceptional. Political murder is rare.
3. There is extensive political imprisonment, or a recent history of such imprisonment. Execution or other political murders and brutality may be common. Unlimited detention, with or without a trial, for political views is accepted.
4. Civil and political rights violations have expanded to large numbers of the population. Murders, disappearances, and torture are a common part of life. In spite of its generality, on this level terror affects those who interest themselves in politics or ideas.

5. Terror has expanded to the whole population. The leaders of these societies place no limits on the means or thoroughness with which they pursue personal or ideological goals". (Standard Codebook, 2023, pp. 1305)

Based on the literature review, I expect that changes in the political terror scale or in the political rights rate affect the number of arrivals in Italy. Moore and Shellman (2004) and Melander and Öberg (2007) reported positive and statistically significant results for the political terror scale variable. Therefore, I expect both variables to be positively correlated to the dependent variable, meaning that an increase in the deterioration of political conditions increase the number of arrivals in Italy. Other variables, such as democratic transition, although included in some prior studies, were not included in my analysis because of almost no variation of this variable in the period of my analysis or due to number of missing observations. Moreover, including more political variables did not increase the R^2 of the model, which is further discussed in the robustness check section.

H4: the deterioration of political conditions in the countries of analysis has a positive impact in the number of arrivals in Italy.

Obstacles or facilitators are measured by (i) the lag of cumulative arrivals in Italy, (ii) women employment rate, and (iii) women political empowerment index. The first variable aims to account for the network effect, meaning that migrants arriving in Italy share information with migrants that want to cross the Mediterranean. The higher the number of migrants arriving in Italy, more information would be available for migrants that intent to cross the Mediterranean.

The other two variables aim to account for facilitators of women migration. Both variables are available at the Quality of Government Dataset and aims to measure women financial independence/resources and gender social constrains. I expect the three variables to be positively associated with arrivals in Italy. Women political empowerment is defined below.

"How politically empowered are women?Clarifications: Women's political empowerment is defined as a process of increasing capacity for women, leading to greater choice, agency, and participation in societal decision-making. It is understood to incorporate three equally-weighted dimensions: fundamental civil liberties, women's open discussion of political issues and participation in civil society organizations, and the descriptive rep- resentation of women in formal political positions. Aggregation: The index is formed by taking the average of women's civil liberties index, women's civil society participation index, and women's political participation index." (Standard Codebook, 2023, pp. 1394)

Finally, control variables are included in the model: time and seasonality (control for year trends and month variation) and population size (to control for countries with a bigger population number proportionally producing more refugees in comparison to countries with a lower population number).

The method used for this analysis is fixed effects regression and the results are presented in percentage change of arrivals in Italy, since the dependent variable is logged. Fixed Effects regressions are used to analyze if changes of conflict, economic conditions, and political factors in the countries of origin had an effect on the number of migrants arriving in Italy. The one-way fixed effect is used to control for time-invariant characteristics of countries, such as location and colonial ties. Time and season trends are controlled by inserting the variable "year" and "month" in the regression equations. Since fixed effects regression measures the within country variation of the dependent variable, it is not suitable for cross sectional analysis.

The analysis is divided in two parts. First, I explore the main causes of migration via Central Mediterranean using all data available and then I use the model to test for gender heterogeneity by using the number of arrivals of migrant women and men in Italy as dependent variables.

Therefore, models 1 and 2 are Fixed Effect regressions in which the dependent variable is the log of monthly arrivals in Italy disaggregated by country of origin. The difference between the two models is only the timeframe. For Model 1, the complete timeframe of 2011 to 2018 is included. The timeframe for Model 2 is 2014 to April 2018, since gender disaggregation of arrivals in Italy is only available for this period.

The second part of the analysis explore the gender heterogeneity in migration behavior. Therefore, models 3 and 4 use monthly arrivals of migrant women and men as the dependent variable, respectively, disaggregated by country of origin. I expect the coefficient of the independent variables to differ from model 3 and 4 in size.

H5: conflict affects the migration behavior of men and women in different intensities.

All models use the same independent and control variables. The variables conflict type and civilian casualties are lagged by 10 months to address the relation of violence and migration timing. Schon (2019) argues that migration timing is not identical to conflict or violence timing because the decision to migrate depends not only on individual motivation but also on opportunity (individual's financial resources, social connection, road safety etc.). The author found a difference of 1 year on average in migration timing between respondents who had higher social resources than respondents with lower resources. Schmeidl (1997) also uses a one-year lag for some variables measuring conflict. Weber (2019), however, uses 3 months lag and argues that a one-year lag in his research showed no result.

I test different lags for the models and find that a lag of 3-to-12-month work for model 1, presenting a statistically significant value for the conflict variables. However, for model 2, the lag of 8-to-12 month works better, with the variable of civilian fatalities being statistically insignificant for lower values of lag. For model 4 (male model) the lag of 5-to-12 month works, but for model 3 (female model) only using the lag 11 and 12 month return statistically significant results. The interpretation is further discussed in the Regression Results part.

Descriptive Results

Figure 2 shows a spaghetti plot of the main countries of origin of the arrivals in Italy from 2011 to 2018. Tunisia, Syria, Eritrea and Nigeria were the countries producing the highest number of migrants crossing the Mediterranean in this analysis. It is possible to notice that the number of arrivals from these countries is much higher than the rest of the sample and each country is responsible for a peak of arrivals of each year of the analysis. From 2014 onwards, there were more countries producing higher numbers of migrants crossing the Mediterranean in comparison to before. In 2010 and 2011, the Tunisian Revolution happened in the country, with a series of protests and confrontations that led the Ben Ali dictator to step out of power. In Syria, a civil war is happening in the country since 2011, whereas in Eritrea the Eritrean-Ethiopian border conflict stated in 1998. In Nigeria, national presidential election was hold in 2015, same year Boko Haram conducted a series of suicide bombings in the country.

Figure 2: Main countries of origin of arrivals in Italy from 2011 to 2018

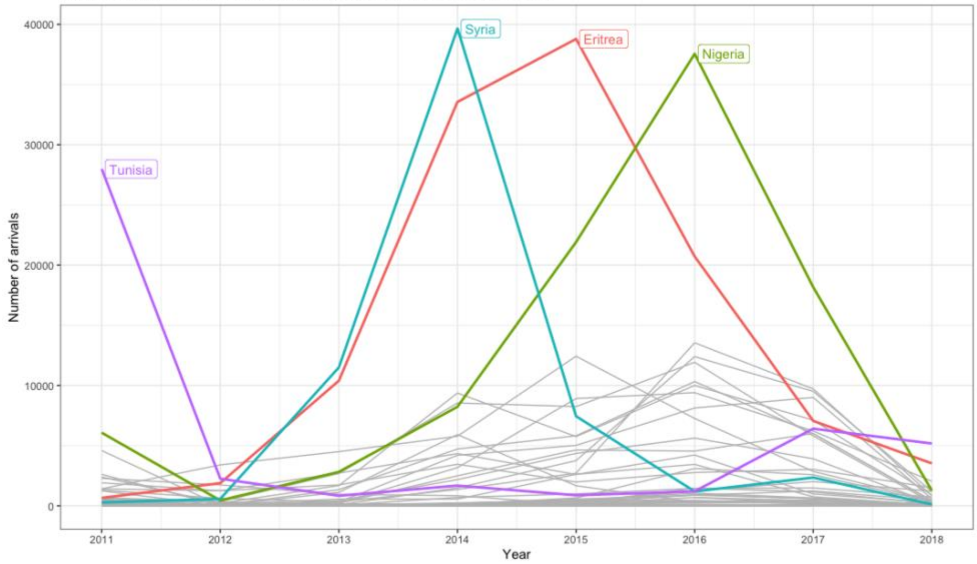


Figure 3 shows the distribution of number of countries in the sample producing migrants versus countries not producing migrants per year. The year of 2015 is the one with the higher number of countries producing migrants in the sample whereas 2012 is the year with the lowest number of countries producing migrants. Besides 2012, the majority of countries in sample produced migrants crossing the Central Mediterranean in the years of analysis.

Figure 3: Countries in the sample producing migrants crossing the Central Mediterranean by year

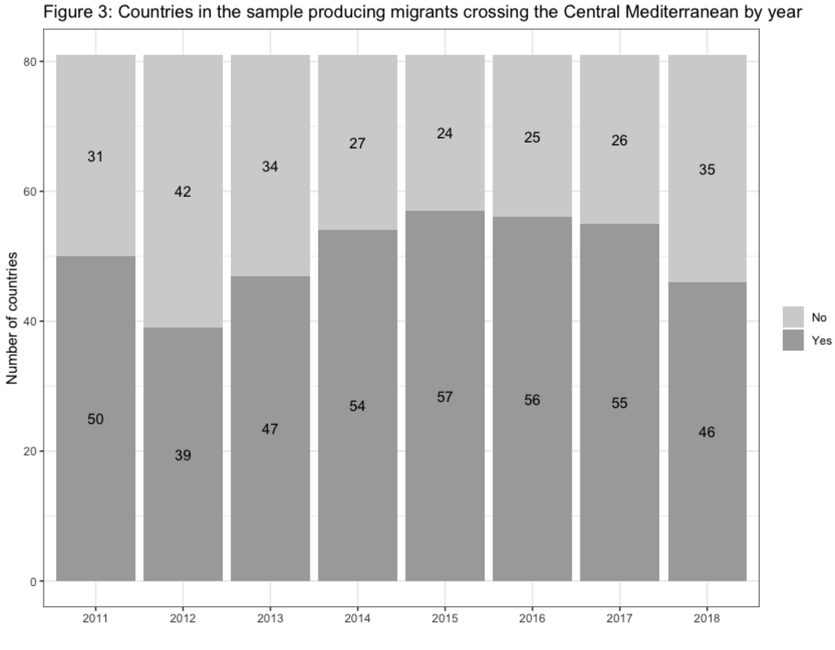


Figure 4 shows the distribution of number of civil deaths in the countries of analysis over the years. Syria was the country with the higher number of civilian deaths in all years of analysis. Afghanistan, Nigeria, Iraq, Central Africa Republic, Yemen and DR Congo are also countries from the sample with high numbers of civilian fatalities. The sample median in close to zero.

Figure 4: Civilian deaths in country of origin of arrivals in Italy

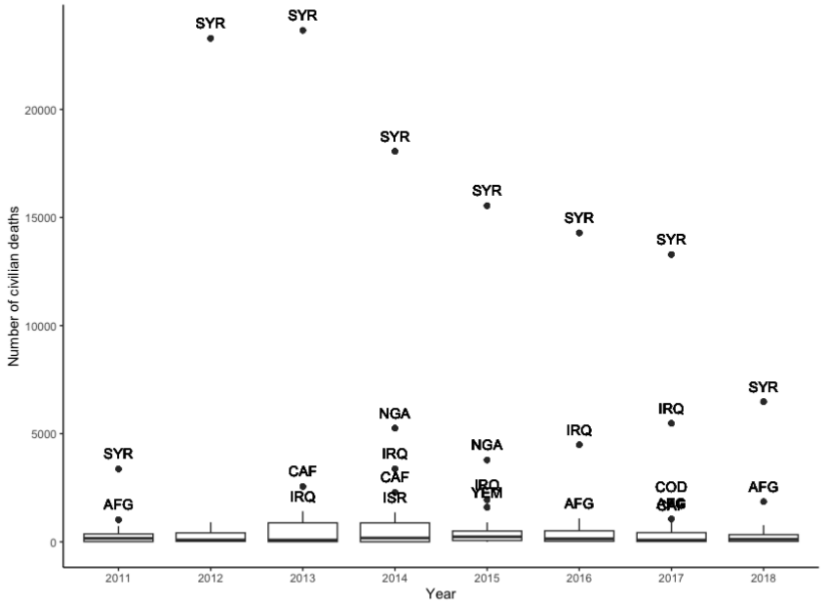
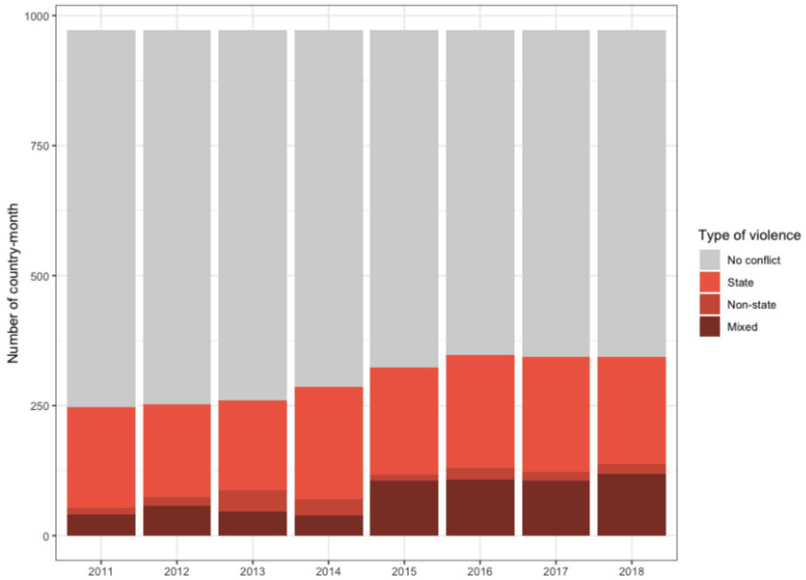


Figure 5 shows the number of country-month conflicts or absence of conflicts. Most of the country-month in the sample were recorded with no conflict. However, there was an increase in number of country-month conflict from 2013 to 2016, with a considerable increase of mixed conflict from 2014 to the subsequent years. Conflicts involving non-state are the ones with the lowest number of records. The dominant type of conflict in the period of analysis was involving States.

Figure 5: Number of country-month by type of conflict



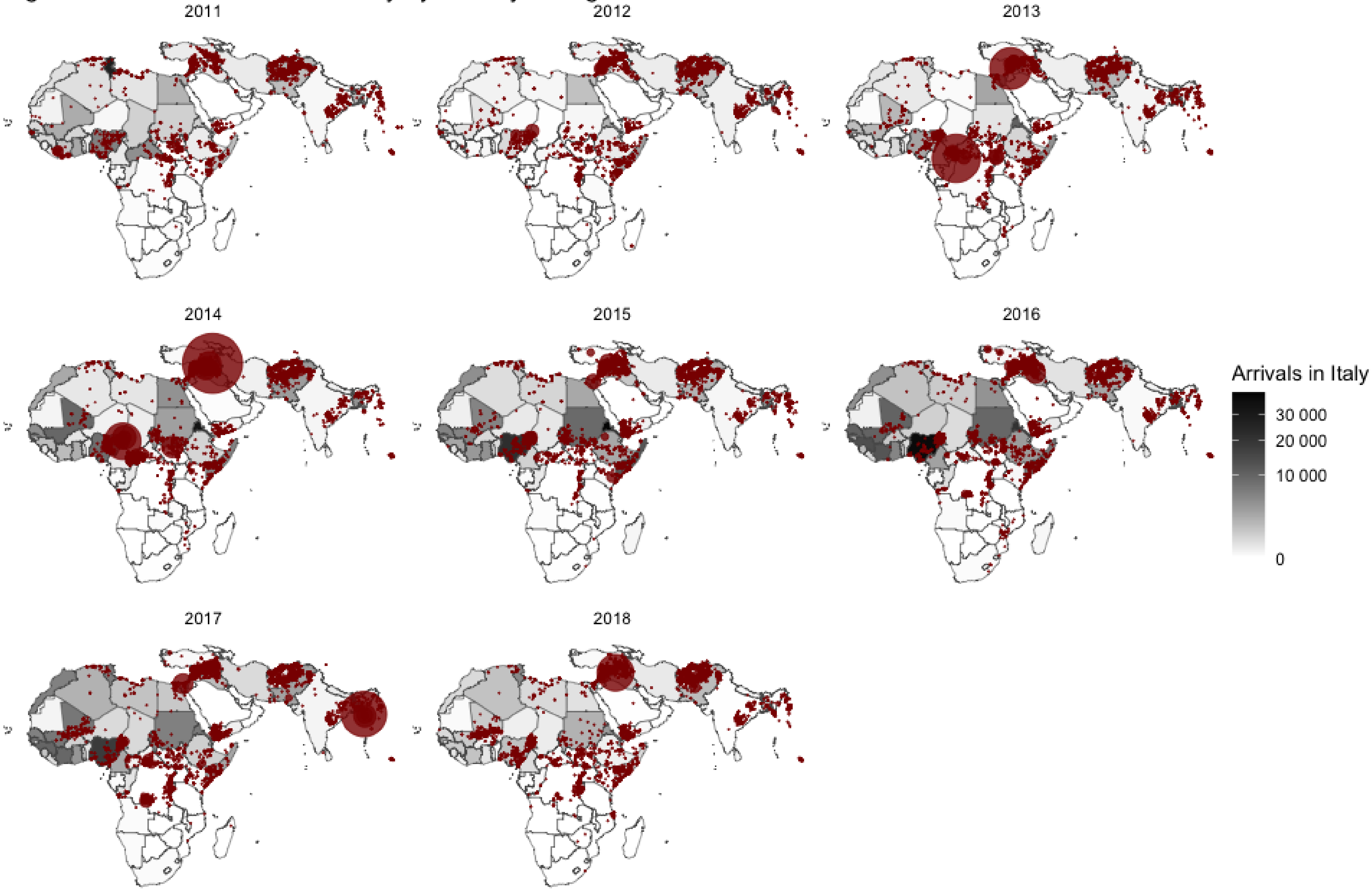
The summary statistics for other variables shows that the mean annual unemployment rate of the sample is 8.58%. The mean and media of women employment rate are low, meaning that countries in the sample have a low percentage of women employed. For political rights, the mean and median are around 5, meaning that most countries in the sample are close to less political freedom category (7).

Table 2: Summary statistics of the independent and control variables

Statistic	N	Min	Max	Median	Mean	St. Dev.
Year	648	2,011	2,018	2,014.5	2,014.50	2.29
Arrival in Italy	648	0	39,651	5	1,107.56	3,894.45
GDP per capita	628	231.45	98,041.36	1,924.61	6,205.31	12,091.48
Unemployment	632	0.11	26.97	6.84	8.58	6.51
Political rights	632	1	7	5	4.67	1.81
Terror scale	632	1	5	3	3.05	1.07
Gender empwr	632	1.55	8.99	6.90	6.36	1.80
Women emplm	624	0.03	10.82	0.88	1.30	1.51
Population	640	87,441	1,369,003,306	11,398,700.0	41,127,687.00	148,481,961.00

Figure 6 shows the main countries of origin of migrants arriving in Italy and the number and intensity of conflicts that happened in their territory, with the intensity calculated by number of civilian deaths. The highest is the red dot in the map, the higher the number of civilian casualties. In years of analysis, Syria, Afghanistan, and Iraq were countries with the greatest number of registered conflicts. Nigeria, DR Congo, Somalia, and Pakistan were also countries in the sample with a high number of conflicts recorded. The years with conflicts with the highest number of civilian casualties happened in 2013 and 2014. Conflict seems to be overall restricted to some areas of the region of analysis. It is also possible to notice variation among main countries producing migrants crossing the Central Mediterranean in the years of analysis.

Figure 6: Number of arrivals in Italy by country of origin and conflict incidence



Regression Results

In Table 3, the results of the models show that one-sided conflict type and civilian casualties are significant and positive related to the dependent variable (number of arrivals in Italy) for models 1, 2 and 4, controlling for changes in economic situation, political factors, facilitators, and time trends. One-side conflict, lagged by 10 months, is not statistically significant for Model Female, but it is statistically significant when using 11 or 12 months as a lag. The one-sided conflict coefficients of the other models are statistically significant using a lag smaller than 10 months. This is evidence that the arrival of women in Italy, in comparison to men, is not only affected in different intensities by one-sided conflicts in the countries of analysis, but also that the migration timing of men and women are different. This result is in accordance with Schon's (2019) findings that men and women have different migration timings.

Table 3: Regression results of Models 1, 2, 3 and 4

	Dependent variable: Arrivals in Italy			
	Model 2011-2018 (1)	Model 2014-2018 (2)	Model Female (3)	Model Male (4)
State-base conflict	-0.036	0.088	-0.031	0.184^{**}
	(0.063)	(0.079)	(0.077)	(0.090)
Non-state conflict	0.312^{***}	0.034	-0.049	-0.006
	(0.111)	(0.149)	(0.145)	(0.170)
One-side conflict	0.545^{***}	0.383^{***}	0.061	0.334^{***}
	(0.075)	(0.095)	(0.092)	(0.109)
Civilian casualties	0.001^{***}	0.001^{**}	0.001^{***}	0.002^{***}
	(0.0001)	(0.0003)	(0.0003)	(0.0003)
Unemployment	-0.133^{***}	-0.032	-0.125^{***}	-0.054
	(0.019)	(0.030)	(0.029)	(0.035)
(Log) GDP per Capita	0.428^{***}	0.858^{***}	1.651^{***}	1.779^{***}
	(0.102)	(0.172)	(0.166)	(0.196)
Political rights	0.019	0.217^{***}	0.172^{***}	0.251^{***}
	(0.035)	(0.059)	(0.057)	(0.067)
Terror scale	0.032	-0.003	-0.078	-0.067
	(0.037)	(0.051)	(0.049)	(0.058)
Women employment rate	-0.402^{***}	0.068	0.008	0.083
	(0.066)	(0.090)	(0.087)	(0.103)
Gender empowerment	0.123^{**}	0.096	0.329^{***}	0.415^{***}
	(0.057)	(0.099)	(0.096)	(0.113)

Network effect	0.00003***	0.00000	0.0001***	0.0001***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
(Log) Population	2.211***	-0.535	1.742	-0.736
	(0.534)	(1.305)	(1.265)	(1.493)
2012	-0.134			
	(0.104)			
2013	0.096			
	(0.106)			
2014	0.634***			
	(0.111)			
2015	0.770***	0.006	0.131	0.098
	(0.116)	(0.103)	(0.100)	(0.118)
2016	1.029***	0.166	0.638***	0.649***
	(0.123)	(0.117)	(0.114)	(0.134)
2017	0.765***	-0.069	0.299**	0.242
	(0.131)	(0.136)	(0.132)	(0.155)
2018	0.190	-0.253	-0.126	-0.291
	(0.152)	(0.165)	(0.160)	(0.188)
February	-0.065	-0.227***	0.038	-0.078
	(0.068)	(0.074)	(0.072)	(0.085)
March	0.140**	-0.053	0.051	0.082
	(0.068)	(0.074)	(0.072)	(0.085)
April	0.372***	0.221***	0.367***	0.451***
	(0.068)	(0.075)	(0.073)	(0.086)
May	0.604***	0.523***	0.488***	0.618***
	(0.071)	(0.082)	(0.080)	(0.094)
June	0.747***	0.550***	0.617***	0.768***
	(0.071)	(0.081)	(0.079)	(0.093)
July	0.767***	0.564***	0.754***	0.936***
	(0.071)	(0.081)	(0.079)	(0.093)
August	0.694***	0.480***	0.783***	0.903***
	(0.071)	(0.080)	(0.077)	(0.091)
September	0.681***	0.237***	0.530***	0.494***
	(0.071)	(0.084)	(0.082)	(0.097)
October	0.596***	0.118	0.423***	0.410***
	(0.071)	(0.084)	(0.082)	(0.097)
November	0.363***	0.010	0.366***	0.335***
	(0.070)	(0.080)	(0.078)	(0.092)

December	0.252^{***}	-0.060	0.378^{***}	0.330^{***}
	(0.070)	(0.080)	(0.078)	(0.092)
Observations	5,966	2,423	2,423	2,423
R^2	0.277	0.149	0.331	0.280
Adjusted R^2	0.264	0.111	0.301	0.248
F Statistic	74.760^{***} (df = 30; 5858)	15.059^{***} (df = 27; 2319)	42.428^{***} (df = 27; 2319)	33.398^{***} (df = 27; 2319)
Note:				*p**p***p<0.01

Table 4 shows the results of Models 3 and 4 using a lag of 11 months for the type of conflict and civilian casualties. The size effect of one-sided conflict on the log of arrivals in Italy is almost twice in the Model Male than in the Model Female. The coefficients of one-sided conflict and civilian casualties are positive and statistic significant at a 1% level. The results show that one-side conflict and civilian casualties are good predictors of arrivals in Italy, confirming the first and second hypothesis of this study. Moreover, it seems that women and men have different migration timing, since the lag of one-sided conflict is statistically significant from a 5 month lag in the male model, and only statistically significant from a 11 month lag for the female model.

Table 4: Regression results of Models 3 and 4, using 11 months lag

	Dependent variable:	
	Arrivals in Italy	
	Model Female	Model Male
	(1)	(2)
State-base conflict	0.006	0.179^{**}
	(0.077)	(0.090)
Non-state conflict	0.087	-0.022
	(0.145)	(0.170)
One-side conflict	0.246^{***}	0.477^{***}
	(0.093)	(0.109)
Civilian casualties	0.002^{***}	0.002^{***}
	(0.0003)	(0.0003)
Unemployment	-0.119^{***}	-0.045
	(0.030)	(0.036)
(Log) GDP per Capita	1.626^{***}	1.798^{***}
	(0.174)	(0.204)
Political rights	0.168^{***}	0.255^{***}
	(0.058)	(0.069)
Terror scale	-0.084^{*}	-0.080
	(0.050)	(0.059)
Women employment rate	-0.005	0.080

	(0.089)	(0.105)
Gender empowerment	0.316***	0.424***
	(0.098)	(0.116)
Network effect	0.0001***	0.0001***
	(0.00000)	(0.00000)
(Log) Population	0.695	-1.995
	(1.335)	(1.570)
Observations	2,346	2,346
R^2	0.339	0.294
Adjusted R^2	0.309	0.261
F Statistic (df = 27; 2242)	42.603***	34.529***
Note:		*p**p***p<0.01

Regarding economic variables, the coefficients for unemployment across models are negatively associated with the dependent variable but are not statistically significant in models 2 and 4. The negative correlation could be explained by the costs of crossing the Mediterranean. The Central Mediterranean route to Italy is expensive and with a higher unemployment rate, it is likely that fewer migrants would have the means to finance the journey. Further investigations are necessary to explain this relationship though. Nonetheless, unemployment rate in the countries of analysis might not be a reliable variable. In many countries in Africa, the informal labour market is a considerable part of the economy, which is not entirely captured by the unemployment rate variable. Moreover, it is difficult to collect reliable information on unemployment rate in conflict zones, like Syria, since the government does not control the entire territory and the level of destruction is high.

The log GDP per capita is positive and statistically significant across all models. This result is in accordance with the literature review that an increase in GDP per capita increases migration because it means that more people have more financial resources to migrate. However, the results of the relation of GDP per capita and the Central Mediterranean migration should be interpreted carefully. There is evidence that the relation might not be linear (Neumayer, 2005) and, as mentioned in the literature review, there are mixed evidence on the effect of GDP on migration. One must be careful when considering economic factors as a root cause of forced migration, which is further discussed in the conclusion.

Regarding political variables, political rights is positive and statistically significant in models 2, 3 and 4; however, the terror scale variable is not statistically significant for any of the models. The political rights variable shows that an increase in the deterioration of political freedom increased arrivals in Italy from 2014 to 2018. The size of the coefficient is higher for the male model than the female one.

The variables used to measure gender equality (as a facilitator of migration) in the countries of analysis show that women employment rate is not a good predictor of arrivals in Italy. The results are not statistically significant for models 2, 3 and 4. The coefficients of Women empowerment index, on the other hand, are positive and statistically significant for models 1, 3 and 4. The size of the coefficients of the female and male models are similar, meaning that this variable affects the migration decision of men and women in similar ways.

The lag of cumulative arrivals in Italy, used to account for network effect, is positive and statistically significant for models 1, 3 and 4. It shows that migrants that crossed the Central Mediterranean to Italy facilitate the journey for other migrants that want to use this route. The mechanism of this relation is the exchange of information between migrants that did the journey and the ones that aim to do the journey. The coefficients are similar in size for the female and male models.

The total sample size comprises of 78 countries and 5,966 observations for Model 1 and 2,423 observations for models 2, 3 and 4. The number of observations is lower because the period of analysis is reduced. A smaller number of observations could affect the significancy of some variables, as it is observed for unemployment, gender empowerment and network effect for model 2. These significant changes from the first to the second model could also indicate a shift in the dynamics of the causes of migration via Central Mediterranean from 2014, or it could be evidence of the effect of exogenous factors on the dependent variable, such as changes in the EU external border control.

The R^2 of the models, as a test of goodness of fit, are around 0.3 for models 1, 3 and 4, meaning that the model predicts around 30% of the variation of the dependent variable. The R^2 of model 2 is lower in comparison to other models, with a decrease of 0.15 points from the first model.

To summarize, one-sided violence and civilian casualties are strong predictors of arrivals in Italy. The results are in line with the literature review that conflict is one of the main causes of forced displacement. Other variables, such as political rights, network effects, women empowerment, and log GDP per capita also seem to be have an effect on the dependent variable.

One-sided conflict and political rights seem to have a different effect size on arrivals of men and women in Italy. On the other hand, network effects, civilian casualties and women empowerment have similar size effects on the arrival of migrant men and women. These results provide evidence that women are differently affected by violence in their migration behavior and timing in comparison to men, therefore hypothesis 5 is accepted.

Although this research provides evidence that causes of flee affect men and women in different intensities and timing, further research is necessary to understand the mechanism behind the gender heterogeneity in migration behavior and timing. It could be that this difference is due to access to resources or by social constrain. Road safety might also play a role in the decision of migrant women to flee, since women are more at risk of suffering gender violence. Or it could be that men migrate earlier to then facilitate the migration of their spouses and children. Unfortunately, the aggregated data of this research does not allow to further explore these mechanisms.

Robustness check and limitations

For the robustness check, I used a correlation table for the independent variables to check for multicollinearity. Only gender empowerment index and political rights were strongly correlated (with a correlation coefficient of 0.7). However, taking them out of the model did not change each other's coefficient much, therefore I decided to leave both variables in the models.

When adding more economic and political variables to the model, there is no change in R^2 of the models and the coefficients of the main independent variables (type of violence and civilian casualties) do not change. The strength of using a Fixed Effect model is that it controls for time-invariant characteristics of countries. However, not including time varying observed or unobserved variables to the model could bias the results. A factor that could cause omitted variable bias is, for example, changes in asylum policies in Europe recognizing refugees from certain nationalities.

Another variable, not easily measurable nor estimated, that could bias the results is the dangerousness of the route. It could be that women face additional risks when trying to reach Europe because of gender based violence or they could be more risk averse than men. It could also be that people from different countries have access to different information on the dangerousness of the route. It is likely that news of dead and missing people in the Mediterranean are not equally spread across all countries of analysis.

Moreover, Melander & Öberg (2006) finds that the geographical scope of conflict affects the intensity of displaced persons. Conflicts happening in urban centers are an important predictor of the variation in the number of forced migrants. The dataset used for conflict does not include information on rural versus urban areas. But differently from Melander & Öberg (2006), I use civil deaths as a measure for conflict intensity, since I assume that it is more likely that indiscriminate violence happening in urban areas would lead to a higher number of civilian casualties. I assume that since there are a higher concentration of people living in urban than rural areas, it is likely that number of civil deaths is partially covering the effect of geographical scope.

Figure 6 shows that some conflicts happen in border regions. It is likely that some conflicts have spillover effects in neighboring countries, which could affect internal and external migration. However, the spillover effect was not accounted in the model for the sake of simplicity. Further research would be necessary to understand if conflict affect internal or external migration of a neighboring countries.

Some other limitations include the data used for the analysis and the chosen methodology. The data used for the dependent variable is arrival data collected by the Italian authorities and compiled by Frontex, IOM and UNHCR. It does not capture all individuals entering in Italy by the Central Mediterranean route as there are several missing information in the dataset. For example, some country of origin in the dataset is coded as "Unknown", "Unspecified Sub-Sahara" and "Stateless". These cases were excluded from the model. For data disaggregated on gender, there are a lot of more missing observations.

There were other independent variables not used in the analysis because of the missing observations in the dataset, such as poverty and homicide rates. When including these variables, I miss more than 50% of my observations.

Finally, although the independent variables are measured on a country-year unit and the dependent variable and conflict variables are measured on a country-month-year unit, using monthly data allows for a more precise measurement of how conflict affects migration timing and flow via Central Mediterranean. Monthly data allows for exploring the different lag between a conflict event and migration timing. I assume that macro-level variables like unemployment rate, GDP per capita, political rights etc. do not present much monthly variation.

Conclusion and policy recommendations

This research aims to present evidence on the main drivers of the migration via Central Mediterranean and how women and men are differently affected by them. The results of this research complement and support the findings of other studies addressing the causes of forced displacement, and more specifically the causes of the 2015 refugee crisis in Europe. The following policy recommendations are an attempt to link empirical evidence to policy making, in order to promote evidence-based policies.

The results show that one-sided conflict and number of civilian fatalities have a significant effect across all regression models of this research. This corroborates previous findings that violence is one of the main causes of migration. Evidence is also found for the effect of political variables (such as political rights), network effects, and women empowerment on the dependent variable. However, the coefficients are not statistically significant across all models.

A positive effect of log GDP per capita on the dependent variable is also found, however, the interpretation of this variable should take careful consideration. First, prior studies present mixed evidence of the direction of the relation of GDP per capita and migration. Some authors find a positive relation, others a negative one or no effect.

Finally, gender heterogeneity is found for one-sided conflict and political rights. The size of the effect of one-sided conflict variable is twice bigger for the male model than the female model. The effect size of the political rights variable is algo higher for the male model. This could either indicate that men are more affected by conflicts and the deterioration of political rights than women, or that financial and social constrains affect women decision to flee even when conflict increases in intensity, or the political situation deteriorates.

Another interesting finding concerning the gender heterogeneity of migration is the migration timing. The results might be evidence that women and men have different migration timings. Men might flee the country first to escape forced military service, or they might migrate to Europe first to facilitate the migration of other family members through family reunification policies.

Further research would be necessary to explore why and how women and men are differently affected in their decision and time to migrate. Ideally, micro-level data should be used to explore individual's motivations.

Nonetheless, the results of this research can be used to draw policy recommendations. First, the results do not corroborate with EU and international development agencies' programs that use development policies to address root causes of migration. The reason and motivation of financing such projects in Africa and Asia are that underdevelopment and poverty are the main drivers of the migration to Europe. However, the results of this research and previous studies show that this might not correspond to the reality. Therefore, it seems unrealistic that development policies promoting economic opportunities or job creation in Africa and Asia are alone going to address the migration crisis to Europe.

Geiger and Pécoud (2013) suggest a critical approach to analyze the migration-development nexus, understanding it as a product of political construction, instead of accepting this relation as given. It is important to understand who are the actors involved in this narrative, and what are their interests and believes. The authors use as an example the imbalance in the migration-development nexus between development as an instrument of fighting the root causes of migration versus migration as freedom of movement of labour force. Europe seems to choose the former approach instead of the later regarding international migration.

Not accepting the migration-development nexus as a natural relation implies in rethinking how development can affect migration (and vice versa) and what empirical evidence is available of what works and what does not work in terms of policy impact. Well-designed development policies must have a clear theory of change, with realistic outputs and outcomes. The problem of development projects tracking root causes of migration is that the causal relationship of development and migration is not as simple as one think and might not lead to the desire outcomes. The results of this research show that an increase in GDP per capita increases the arrivals in Italy via Central Mediterranean, instead of having the opposite effect.

This does not mean that I do not support development projects in Africa, or that development projects addressing the root causes of migration should be suspended. It just means that the justification and the overall logic of the project might not correspond to the reality, which could lead to political frustration or to a feeling that the development policies implemented in Africa are not achieving its goals.

Therefore, my policy recommendation is to link a peace component to the development approach to address root causes of migration. For example, the humanitarian-development-peace (HDP) approach seeks to integrate the work of humanitarian, development and peace organizations to provide joint responses to crisis, instead of using a divided approach (Kittaneh & Stolk, 2018; Centre on International Cooperation, 2019). Humanitarian approaches in protected crisis often perpetuate the need of international aid, for national structures are not equipped to address the crisis alone. On the other hand, development approaches are costly, the results are not immediate and, as shown in the results of this research, are not sufficient to address forced displacement. Therefore, the peace component is indispensable to reach sustainable peace and long-term development.

My second policy recommendation is to include a gender component in migration policies and programs and to promote more research in this field. This research provides empirical evidence that women and men are differently affected in their migration patterns. Other authors explore how migration laws and policies create different patterns of migration among men and women. Therefore, it is important to study how certain projects and policies in the field of migration affect women, and how these projects can create negative effects or unintended results, as illustrated by Gerard and Pickering (2013).

These policies recommendations are an effort to link evidence and policy to produce more evidence-based policies in the field of migration.

Bibliography

Adhikari, P. (2013). Conflict-induced displacement, understanding the causes of flight. American Journal of Political Science, 57(1), 82-89.

Aksoy, C., & Poutvaara P. (2019). Refugees' Self-selection into Europe: Who Migrates Where? Leibniz Institute for Economic Research at the University of Munich, ifo Working paper, 289, 1-26.

Blangiardo, G. (2012). Gender and migration in Southern and Eastern Mediterranean and Sub-Saharan African countries. CARIM Research Report 2012/01, Robert Schuman Centre for Advanced Studies, San Domenico di Fiesole (FI): European University Institute.

Bohra-Mishra, P., & Massey, D. S. (2011). Individual decisions to migrate during civil conflict. Demography, 48(2), 401-424.

Brück, T. & Dunker, K. & Ferguson, N. & Meysonnat, A. & Nillesen, E. (2018). Determinants and Dynamics of Forced Migration to Europe: Evidence from a 3-D Model of Flows and Stocks. IZA Discussion Paper, 11834, 1-33.

Center on International Cooperation (2019). The Triple Nexus in Practice: Towards a New Way of Working in Protracted and Repeated Crisis. New York: Center on International Cooperation.

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

Geiger, M. & Pécoud, A. (2013). Migration, Development and the 'Migration and Development Nexus'. Population, Space and Place, 19(4), 369-374.

Gerard, A. & Pickering, S. (2013). Gender, Securitization and Transit: Refugee Women and the Journey to the EU. Journal of Refugee Studies, 27(3), 338-359.

Weber, H. (2019). Can Violent Conflicts Explain the Recent Increase in the Flow of Asylum Seekers From Africa Into Europe? Journal of Immigrant & Refugee Studies, 17:4, 405-424, DOI: 10.1080/15562948.2018.1517424

Kirwin, M. and J. Anderson (2018). “Identifying the factors driving West African migration”. OECD Publishing West African Papers, 17, 1-19.

Massey, D. et al. (1993). Theories of international migration: A review and appraisal. Population and Development Review, 19, 431-466.

Melander, E., & Öberg, M. (2006) Time to Go? Duration Dependence in Forced Migration, International Interactions, 32(2), 129-152, DOI: 10.1080/03050620600574873

Melander, E., & Öberg, M. (2007). The Threat of Violence and Forced Migration: Geographical Scope Trumps Intensity of Fighting. Civil Wars, 9(2), 156-173, DOI: 10.1080/13698240701207310

Moore, W., & Shellman, S. (2007). Whither will they go? A global study of refugees’ destinations, 1965–1995. International Studies Quarterly, 51(4), 811-834.

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

Müller-Funk, L. (2019). Adapting to staying, or imagining futures elsewhere: Migration decision-making of Syrian refugees in Turkey. International Migration Institute, Working Papers, No. 154.

Neumayer, E. (2005). Bogus Refugees? The Determinants of Asylum Migration to Western Europe. International Studies Quarterly, 49(3), 389–409. http://www.jstor.org/stable/3693600

Niemann, A. & Zaun, N. (2018). EU Refugee Policies and Politics in Times of Crisis: Theoretical and Empirical Pespectives. Introduction to the Special Issue, Journal of Common Market Studies, 56(1), 3-22.

Piper, N. (2006). Gendering the Politics of Migration. The International Migration Review,Gender and Migration Revisited, 40(10), 133-164.

Schmeidl, S. (1997). Exploring the causes of forced migration: A pooled time-series analysis, 1971-1990. Social Science Quarterly, 78(2), 284-308.

Schon, J. (2019). Motivation and opportunity for conflict-induced migration: An analysis of Syrian migration timing. Journal of Peace Research, 56(1), 12-27.

Kang, Y. (2021). Refugee crisis in Europe: determinants of asylum seeking in European countries from 2008–2014. Journal of European Integration, 43(1), 33-48.

Kittaneh, A. & Stolk, A. (2018). Doing Nexus Differently: How can Humanitarian and Development Actors link or integrate humanitarian action, development, and peace? Research Paper, CARE Regional Applied Economic Empowerment Hub.

Data Reference

Croicu, M. & Sundberg, R. (2016). UCDP GED Codebook version 5.0. Department of Peace and Conflict Research, Uppsala University.

Davies, S.; Pettersson, T.; Öberg, M. (2022). Organized violence 1989-2021 and drone warfare. Journal of Peace Research 59(4).

World Bank. World Bank Open Data. Retrieved from <https://data.worldbank.org>

Pemstein, D.; Marquardt, K.; Tzelgov E.; Wang, Y.; Medzihorsky, J.; Krusell, J.; Miri, F.; Römer, J. (2022). The {V-Dem} Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data. University of Gothenburg, Varieties of Democracy Institute: Working Paper No. 21, 7th edition.

International Organization for Migration. Displacement Tracking Matrix Datasets. IOM< Global Data Institute. Retrieved from <https://dtm.iom.int>

Sundberg, Ralph and Erik Melander (2013). Introducing the UCDP Georeferenced Event Dataset. Journal of Peace Research 50(4).

Teorell, Jan, Aksel Sundström, Sören Holmberg, Bo Rothstein, Natalia Alvarado Pachon, Cem Mert Dalli & Yente Meijers (2023). The Quality of Government Standard Dataset, version Jan23. University of Gothenburg: The Quality of Government Institute, https://www.gu.se/en/quality-government doi:10.18157/qogstdjan23

United Nations High Commissioner for Refugees. Refugee Situations: Mediterranean Situation. IOM, Operational Data Portal. Retrieved from https://data.unhcr.org/en/situations/mediterranean

R package citation

Croissant Y, Millo G (2008). "Panel Data Econometrics in R: The plm Package." _Journal of Statistical Software_, *27*(2), 1-43. doi:10.18637/jss.v027.i02 <https://doi.org/10.18637/jss.v027.i02>.

Hlavac, Marek (2022). stargazer: Well-Formatted Regression and Summary Statistics Tables. R package version 5.2.3. https://CRAN.R-project.org/package=stargazer

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms

J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, WooK, Yutani H (2019). “Welcome to the tidyverse.” _Journal of Open Source Software_, *4*(43), 1686. doi:10.21105/joss.01686 <https://doi.org/10.21105/joss.01686>.

Wickham H, Miller E, Smith D (2022). _haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files_. R package version 2.5.1, <https://CRAN.R-project.org/package=haven>.

Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

Appendix - R code

```
## Data

#Frontex dataset
detection_data <- read_excel("detection_data.xlsx")

#Conflict data
conflict_death <- read_excel("GEDEvent_v22_1-2.xlsx")

#ISO list
ISO <- read_excel("ISO_complete.xlsx")

#GDP per capita
GDP_per_capita <- read_excel("GDP_per_capita.xls")

#Unemployment
unemployment <- read_excel("unemployment.xls")

#Population size
population <- read_excel("population.xls")

#political variables

political <- read_csv("qog_std_ts_jan23.csv")

#2014 arrival gender
data_2014 <- read_excel("data_2014.xlsx")

#2015 arrival gender
data_2015 <- read_excel("data_2015.xlsx")

#2016 arrival gender
data_2016 <- read_excel("data_2016.xlsx")

#2017 arrival gender
data_2017 <- data_2017 <- read_excel("data_2017.xlsx")
```

```
#2018 arrival gender
data_2018 <- read_excel("data_2018.xlsx")

#data transformation

#detection data
detection_data <- pivot_longer(data = detection_data, cols = 4:135,
                  names_to = c("month", "year"),
                  names_pattern = "([A-Za-z]+)(\\d+)",
                  values_to = "arrival")

detection_data <- rename(detection_data, country= nationality)

detection_data <- mutate(detection_data , month_number = case_when(month ==
"JAN" ~ 1,
                    month == "FEB" ~ 2,
                    month == "MAR" ~ 3,
                    month == "APR" ~ 4,
                    month == "MAY" ~ 5,
                    month == "JUN" ~ 6,
                    month == "JUL" ~ 7,
                    month == "AUG" ~ 8,
                    month == "SEP" ~ 9,
                    month == "OCT" ~ 10,
                    month == "NOV" ~ 11,
                    month == "DEC" ~ 12))

detection_data$year  <- as.numeric(detection_data$year)

detection_data$date <- paste(detection_data$year, detection_data$month_number,
sep="-") %>% ym() %>% as.Date()


#Conflict data
conflict_death <- filter(conflict_death, year >= 2011 & year <= 2018)

conflict_death$date_conflict <- format(as.Date(conflict_death$date_start), "%Y-%m-
%d")
conflict_death$month_number <- format(as.Date(conflict_death$date_start), "%m")
conflict_death$month_number  <- as.numeric(conflict_death$month_number)

colnames(conflict_death)

conflict_death <- select(conflict_death, year, type_of_violence, country, region,
deaths_civilians, date_conflict,month_number, latitude, longitude)

conflict_death <- conflict_death %>% group_by(country,month_number, year) %>%
  mutate(total_death_month = sum(deaths_civilians))
```

```
conflict_death$date_conflict <- as.Date(conflict_death$date_conflict, "%Y-%m-%d")
class(conflict_death$date_conflict)

table(conflict_death$country)

conflict_death$country[conflict_death$country == "DR Congo (Zaire)"] <- "DR Congo"
conflict_death$country[conflict_death$country == "Madagascar (Malagasy)"] <- "Madagascar"
conflict_death$country[conflict_death$country == "Yemen (North Yemen)"] <- "Yemen"
conflict_death$country[conflict_death$country == "Zimbabwe (Rhodesia)"] <- "Zimbabwe"


conflict_death$conflict <- 1

#GDP per capita
GDP_per_capita <- pivot_longer(data = GDP_per_capita, cols = 2:63,
                               names_to = "year",
                               values_to = "GDP_per_capita")
GDP_per_capita$year  <- as.numeric(GDP_per_capita$year)

#GDP_per_capita$GDP_per_capita  <-format(round(GDP_per_capita$GDP_per_capita, 2), nsmall = 2)

#Unemployment
unemployment <- pivot_longer(data = unemployment, cols = 2:63,
                             names_to = "year",
                             values_to = "unemployment")
unemployment$year  <- as.numeric(unemployment$year)

#Population size
population <- pivot_longer(data = population, cols = 2:63,
                           names_to = "year",
                           values_to = "population")
population$year  <- as.numeric(population$year)

population <- rename(population, ISO = `Country Code`)

#ISO
ISO <- select(ISO, c("country", "ISO", "region", "sub-region"))

#political factors

political <- filter(political, year >= 2011 & year <= 2018)

political <- rename(political, ISO = ccodealp)
```

```
political$ISO[political$ISO == "SSD"] <- "SDS"

#merge datasets

data <- left_join(detection_data, conflict_death, by = c("year", "month_number",
"country"))

data <- left_join(data, ISO, by = "country")

data <- left_join(data, GDP_per_capita, by = c("ISO", "year"))

data <- left_join(data, unemployment, by = c("ISO", "year"))

data <- left_join(data, population, by = c("ISO", "year"))

data <- left_join(data, regime_change, by = c("ISO", "year"))

data <- left_join(data, political, by = c("ISO", "year"))

# Create new variables and remove duplicates

data$deaths_civilians[is.na(data$deaths_civilians)] <- 0

data$conflict[is.na(data$conflict)] <- 0

#remove duplicates (when conflict happened twice in a month)

data <- data %>% distinct(country,month,year,arrival, route, .keep_all = TRUE)

data$total_death_month[is.na(data$total_death_month)] <- 0

data$type_of_violence[is.na(data$type_of_violence)] <- 0

## Plot total arrivals in Italy and Spain

data %>% group_by(route,year, month_number) %>%
  summarise(total = sum(arrival))%>%
  ggplot(aes(x= as.factor(month_number), y = total)) +
  geom_col(fill="grey") +
  facet_wrap(~route+year)+
  ggtitle("Figure 1: Monthly arrivals via Central and Eastern Mediterranean Routes from
2011 to 2018") +
  ylab("number arrivals")+
  xlab("month") +
  theme_classic()

#restrict analysis to 2011-2018
```

```
data <- filter(data, year >= 2011 & year <= 2018)

#exclude countries not relevant to the analysis: in America, Europe, Oceania, Asia etc
data <- data %>% filter(country != "Unknown") %>%
  filter(country != "Albania") %>%
  filter(country != "Belarus") %>%
  filter(country != "Colombia") %>%
  filter(country != "French Guiana") %>%
  filter(country != "Kosovo") %>%
  filter(country != "Moldova") %>%
  filter(country != "Panama") %>%
  filter(country != "Russia") %>%
  filter(country != "Stateless") %>%
  filter(country != "Ukraine") %>%
  filter(country != "China")%>%
  filter(country != "Haiti")%>%
  filter(country != "Kyrgyzstan")%>%
  filter(country != "Malaysia")%>%
  filter(country != "Myanmar")%>%
  filter(country != "Tajikistan")%>%
  filter(country != "Uzbekistan")%>%
  filter(country != "Vietnam") %>%
  filter(country != "Belize")%>%
  filter(country != "Bermuda")%>%
  filter(country != "Sub-Sahara")%>% ### not sure what to do with this
  filter(country != "Kazakhstan")%>%
  filter(country != "Jamaica")

#exclude more countries based on sub-region

data <- rename(data,sub_region = `sub-region`)
data <- data %>% filter(!sub_region %in% c("Latin America and the Caribbean",
"Southern Europe", "South-eastern Asia", "Eastern Asia"))

## Subset data to Central Mediterranean only
data_c <- data %>% filter (route == "Central Mediterranean Route")

## Data visualization -Dependent variable

data_2 <- data_c %>% filter(arrival > 1) %>% group_by(country, year) %>%
summarise(total_arrival = sum(arrival)) %>%
  filter(country %in% c("Syria", "Nigeria", "Eritrea", "Tunisia"))

data_label <- data_2 %>% group_by(country) %>% filter(country %in% c("Syria",
"Nigeria", "Eritrea", "Tunisia")) %>%
  group_by(country) %>%
  arrange(-total_arrival) %>%
  top_n(1)
```

```
data_c %>% filter(arrival > 1) %>% group_by(country, year) %>%
summarise(total_arrival = sum(arrival)) %>%
  ggplot(aes(x = year, y = total_arrival, group = country)) +
  scale_x_continuous(breaks = seq(2011, 2018, 1)) +
  geom_line(color= "grey")+
  geom_line(data = data_2, aes(x = year, y = total_arrival, color = country), size = 1) +
  geom_label(data = data_label,
           aes(x = year + 0.03, label=country, color = country), hjust=0)+
  theme(
       legend.position = "none")+
  ylab("Number of arrivals")+
  xlab("Year")+
  ggtitle("Main countries of nationality of arrivals in Italy")

#Plot Countries producing migrants vs countries not producing migrants

Fig_2 <- Fig_2 %>% mutate(dummy = case_when(total >= 1 ~ 1,
                                    total == 0 ~ 0))

Fig_2 %>% group_by(dummy, year) %>% summarise(n = n()) %>%
  ggplot(aes(x = as.factor(year), y = n, fill = as.factor(dummy)))+
          geom_bar(stat = "identity")+
  geom_text(aes(label = n), position = position_stack(vjust = 0.5))+
  scale_fill_manual(name = "", labels = c("No", "Yes"),
                values = c("lightgrey", "darkgrey"))+
  ggtitle("Figure 3: Countries in the sample producing migrants crossing the Central
Mediterranean by year")+
  ylab("Number of countries")+
  xlab("")+
  theme_bw()

#Intensity plot

data_c %>% select(year, month_number, type_of_violence, total_death_month, ISO)
%>%
  filter(type_of_violence != "0") %>% group_by(ISO, year) %>%
  mutate(total_death_year = sum(total_death_month)) %>%
  ggplot(aes(x = as.factor(year), y = total_death_year)) +
  geom_boxplot()+
  geom_text(aes(label = ISO), data = data_c %>% filter(type_of_violence != "0") %>%
group_by(ISO, year) %>%
           mutate(total_death_year = sum(total_death_month)) %>%
filter(total_death_year > 1000), vjust= -1)+
  ylab("Number of civilian deaths")+
  xlab("Year")+
  ggtitle("Civilian deaths in country of origin of arrivals in Italy")+
  theme_classic()
```

#Conflict type plot

```
data_c %>% group_by(type_of_violence, year) %>% summarise(n = n()) %>%
  ggplot(aes(x = as.factor(year), y = n, fill = as.factor(type_of_violence)))+
  geom_bar(stat = "identity")+
  scale_fill_manual(name = "Type of violence", labels = c("No conflict", "State", "Non-state", "Mixed"),
                    values = c("lightgrey", "coral2", "coral3", "coral4"))+
  ggtitle("Number of country-month by type of conflict")+
  ylab("Number of country-month")+
  xlab("Year")+
  theme_bw()
```

#Summary statistics

```
data_c %>% group_by(ISO, year) %>% mutate(total_arrival = sum(arrival)) %>% filter(month_number ==1) %>%
  select(ISO, total_arrival, GDP_per_capita, unemployment,fh_pr, gd_ptss, vdem_gender, wdi_empf,population,)%>%
  as.data.frame() %>%
  stargazer(type = "text", #or html or latex
          out = "html_table.html",
          digits=2,
          summary.stat = c("n", "min", "max", "median", "mean", "sd"))
```

Geograpgical

```{r}
Fig_2 <- data_c %>% group_by(year, ISO) %>%
  summarise(total = sum(arrival))

Fig_2$year <- as.factor(Fig_2$year)

Fig_2 <- rename(Fig_2, adm0_a3 = ISO)

conflict_death <- left_join(conflict_death, ISO, by = "country")

table(conflict_death$region.x)

Fig_conflict <- conflict_death %>% filter(region.x == "Asia" | region.x == "Africa" | region.x == "Middle East" )

Conflict <- Fig_conflict %>% group_by(year, country) %>% summarise(n())

world <- ne_countries(scale = "medium", returnclass = "sf")
class(world)
```

```
ggplot(data = world) +
  geom_sf()

map <-
  ne_countries(scale = "medium", returnclass = "sf")  %>%
  left_join(Fig_2 ,
          by = "adm0_a3") %>%
  filter(year == "2011" | year == "2012" |  year == "2013" | year == "2014" |year == "2015"
| year == "2016" |year == "2017" | year == "2018")

fig_map <- ggplot(data = map) +
  geom_sf(aes(fill = total)) +
  geom_point(data = Fig_conflict, aes(x = longitude, y = latitude), size =
Fig_conflict$deaths_civilians/70, color = "darkred", alpha = 0.8)+
  coord_sf(xlim = c(-25, 105), ylim = c(50, -40), expand = FALSE)+
  scale_fill_gradient(low = "white", high = "black", na.value = "grey80", trans = "sqrt",
                   labels = scales::number_format(accuracy = 1)) +
  facet_wrap(~year)+
 theme(legend.position = "bottom",
        legend.box = "horizontal")+
  xlab("") + ylab("") +
  labs(fill = "Arrivals in Italy")+
  ggtitle("Figure 6: Number of arrivals in Italy by country of origin and conflict
incidence")+
    theme_void()

## Model for the whole period 2011 to 2018

data_c <- filter(data_c, year >= 2011)
data_c$vdem_gender <- data_c$vdem_gender*10
data_c$type_of_violence <- as.factor(data_c$type_of_violence)
data_c <- data_c %>%
  group_by(ISO) %>%
  mutate(cumulative_arrival = cumsum(arrival))

model03b <- plm(log(arrival+1) ~ lag(type_of_violence,10) + lag(total_death_month,
10)+ unemployment +log(GDP_per_capita) + fh_pr+ gd_ptss+ wdi_empf+
vdem_gender + lag(cumulative_arrival) +log(population)+as.factor(year) +
as.factor(month_number) , data = data_c, model= "within", index = "ISO")

cor(correlation[, c("total_death_month", "unemployment",
            "GDP_per_capita", "population", "vdem_gender", "fh_pr",
            "gd_ptss", "wdi_empf")], use = "complete.obs")

## Model for period 2014 to 2018 and gender models

#2014 arrival data
```

```
colnames(data_2014)

data_2014 <- rename(data_2014, male_total = Arrivals_male)
data_2014 <- rename(data_2014, female_total = Arrivals_female)
data_2014 <- mutate(data_2014 , month_number = case_when(month == "january" ~ 1,
                                    month == "february" ~ 2,
                                    month == "march" ~ 3,
                                    month == "april" ~ 4,
                                    month == "may" ~ 5,
                                    month == "june" ~ 6,
                                    month == "july" ~ 7,
                                    month == "august" ~ 8,
                                    month == "september" ~ 9,
                                    month == "october" ~ 10,
                                    month == "november" ~ 11,
                                    month == "december" ~ 12))

data_2014 <- select(data_2014, year, ISO, male_total, female_total, month_number)
#2015 arrival data

data_2015$Reported_date <- format(as.Date(data_2015$Reported_date), "%m") #add
number of month
data_2015 <- mutate(data_2015, year= 2015) #create column year
data_2015 <- rename(data_2015, month_number = Reported_date) #rename month
column
data_2015$month_number <- as.numeric(data_2015$month_number)
data_2015<-data_2015[order(data_2015$ISO, data_2015$month_number),] #order data
to calculate arrivals per month

data_2015<-data_2015 %>%
  group_by(ISO) %>% mutate(male_total = c(NA, diff(Arrivals_cumulative_male)))

data_2015 <-data_2015 %>%
  mutate(male_total = ifelse(month_number == 1, Arrivals_cumulative_male,
male_total))

data_2015<-data_2015 %>%
  group_by(ISO) %>% mutate(female_total = c(NA, diff(Arrivals_cumulative_female)))

data_2015 <-data_2015 %>%
  mutate(female_total = ifelse(month_number == 1, Arrivals_cumulative_female,
female_total))

#2016 arrival data
#note: this dataset has many problems. Different countries have different number of
observation + many missing values
```

```
data_2016$Reported_date <- format(as.Date(data_2016$Reported_date), "%m") #add number of month
data_2016 <- mutate(data_2016, year= 2016) #create column year
data_2016 <- rename(data_2016, month_number = Reported_date) #rename month column
data_2016 <- rename(data_2016, ISO = ISO_3_origin) #rename ISO column
data_2016<-data_2016[order(data_2016$ISO, data_2016$month_number),] #order data to calculate arrivals per month

data_2016<-data_2016 %>%
  group_by(ISO) %>% mutate(male_total = c(NA, diff(Arrivals_cumulative_male)))

data_2016 <-data_2016 %>%
  mutate(male_total = ifelse(month_number == 1, male_total, Arrivals_cumulative_male))

data_2016<-data_2016 %>%
  group_by(ISO) %>% mutate(female_total = c(NA, diff(Arrivals_cumulative_female)))

data_2016 <-data_2016 %>%
  mutate(female_total = ifelse(month_number == 1, female_total, Arrivals_cumulative_female))

#add 2017 data
#note: also many problems with data, but not as serious as 2016


data_2017$Reported_date <- format(as.Date(data_2017$Reported_date), "%m") #add number of month
data_2017 <- mutate(data_2017, year= 2017) #create column year
data_2017 <- rename(data_2017, month_number = Reported_date) #rename month column
data_2017<-data_2017[order(data_2017$ISO, data_2017$month_number),] #order data to calculate arrivals per month

data_2017<-data_2017 %>%
  group_by(ISO) %>% mutate(male_total = c(NA, diff(Arrivals_cumulative_male)))

data_2017 <-data_2017 %>%
  mutate(male_total = ifelse(month_number == 1, male_total,Arrivals_cumulative_male))

data_2017<-data_2017 %>%
  group_by(ISO) %>% mutate(female_total = c(NA, diff(Arrivals_cumulative_female)))

data_2017 <-data_2017 %>%
  mutate(female_total = ifelse(month_number == 1, female_total, Arrivals_cumulative_female))
```

```
colnames(data_2017)

data_2017 <- rename(data_2017, Arrivals_cumulative = Cumulative_arrivals)

#add 2018 data
data_2018 <- mutate(data_2018, year= 2018) #create column year

data_2018<-data_2018 %>%
  group_by(ISO) %>% mutate(male_total = c(NA, diff(Arrivals_cumulative_male)))

data_2018 <-data_2018 %>%
  mutate(male_total = ifelse(month_number == 1, Arrivals_cumulative_male,
male_total))

data_2018<-data_2018 %>%
  group_by(ISO) %>% mutate(female_total = c(NA, diff(Arrivals_cumulative_female)))

data_2018 <-data_2018 %>%
  mutate(female_total = ifelse(month_number == 1, Arrivals_cumulative_female,
female_total))
colnames(data_2017)

data_2015$month_number <- as.numeric(data_2015$month_number)
data_2016$month_number <- as.numeric(data_2016$month_number)
data_2017$month_number <- as.numeric(data_2017$month_number)

data_2015 <- select(data_2015,
month_number,ISO,Arrivals_cumulative,Arrivals_cumulative_male,
Arrivals_cumulative_female, year, male_total, female_total)

data_2016 <- select(data_2016,
month_number,ISO,Arrivals_cumulative,Arrivals_cumulative_male,
Arrivals_cumulative_female, year, male_total, female_total)

data_2017 <- select(data_2017,
month_number,ISO,Arrivals_cumulative,Arrivals_cumulative_male,
Arrivals_cumulative_female, year, male_total, female_total)

data_2018 <- select(data_2018,
month_number,ISO,Arrivals_cumulative,Arrivals_cumulative_male,
Arrivals_cumulative_female, year, male_total, female_total)

data_gender <- bind_rows(data_2014, data_2015, data_2016, data_2017, data_2018)

data_c_2 <- filter(data_c, year >= 2014 & date <= "2018-04-01")

data_c_2 <- left_join(data_c_2, data_gender, by = c("ISO", "year", "month_number"))
```

```
data_c_2 <- filter(data_c_2, !is.na(male_total))
data_c_2 <- filter(data_c_2, !is.na(female_total))

data_c_2 <- mutate(data_c_2, male_total = ifelse(arrival == 0, 0, male_total))

data_c_2 <- mutate(data_c_2, female_total = ifelse(arrival == 0, 0, female_total))

data_c_2$male_total <- abs(data_c_2$male_total)
data_c_2$female_total <- abs(data_c_2$female_total)

data_c_2 <- data_c_2 %>%
  group_by(ISO) %>%
  mutate(cumulative_arrival_male = cumsum(male_total))

data_c_2 <- data_c_2 %>%
  group_by(ISO) %>%
  mutate(cumulative_arrival_female = cumsum(female_total))

model04b <- plm(log(arrival+1) ~lag(type_of_violence,10) + lag(total_death_month,10)+ unemployment +log(GDP_per_capita) + fh_pr+ gd_ptss+ wdi_empf+ vdem_gender + lag(cumulative_arrival) +log(population)+ as.factor(year) + as.factor(month_number) , data = data_c_2, model= "within", index = "ISO")

model_female_b <- plm(log(female_total+1) ~ lag(type_of_violence,10) + lag(total_death_month, 10)+ unemployment +log(GDP_per_capita) + fh_pr+ gd_ptss+ wdi_empf+ vdem_gender + lag(cumulative_arrival) +log(population) + as.factor(year) + as.factor(month_number), data = data_c_2, model= "within", index = "ISO")

model_male_b <- plm(log(male_total+1) ~ lag(type_of_violence,10) + lag(total_death_month, 10)+ unemployment +log(GDP_per_capita) + fh_pr+ gd_ptss+ wdi_empf+ vdem_gender + lag(cumulative_arrival) +log(population)+ as.factor(year) + as.factor(month_number), data = data_c_2, model= "within", index = "ISO")

stargazer(model03b, model04b, model_female_b, model_male_b, type= "html", covariate.labels = c( "State-base conflict", "Non-state conflict",  "One-side conflict", "Civilian casualties", "Unemployment", "(Log) GDP per Capita",  "Political rights","Terror scale","Women employment rate", "Gender empowerment", "Network effect", "(Log) Population",  "2012", "2013", "2014", "2015", "2016", "2017", "2018", "February", "March","April", "May", "June", "July", "August", "September", "October", "November", "December"), title = "Table 2: Regression results of Models 1, 2, 3 and 4", column.labels = c("Model 2011-2018", "Model 2014-2018", "Model Female", "Model Male"), align=TRUE, dep.var.labels = c("Arrivals in Italy","", "", ""),
        out = "Model1.html")
```

Statement of Authorship