Van Veen Violetta (2062051) and
Vulcano Federica (2039157)

Replication paper: Hatton (2016). 60 Million Refugees: Refugees, Asylum Seekers, and Policy in OECD Countries

Introduction

The number of refugees rose by more than 100% from the beginning of this century, reaching a peak of almost 60 million in 2015. This figure highlights that migration is an increasingly **urgent** issue for policymakers and citizens. The number of refugees also includes asylum seekers, on which the research conducted by Hatton (2016) develops.

The analysis made in 'Refugees, Asylum Seekers, and Policy in OECD Countries' comes as a response to the **research question** which investigates the reasons driving the asylum seekers away from their origin countries and the effects of the asylum policies implemented in the destination countries. It is concluded with some policy advice. There are multiple reasons that lead to seeking asylum. Both the findings of Davenport, Moore, and Poe (2003) and those of Moore and Shellman (2007) claim that genocide, civil wars, dissident conflicts and political regime transactions explain the reasons driving the majority of refugees. Economic conditions in the country of origin, expressed as GDP per capita, also influence asylum flows, together with the life prospects in the destination countries. As for the tough policies implemented by the destination countries to decrease the number of asylum applicants, the situation is more complicated. These policies make the process harder and therefore act as a deterrent, however people escaping war, violence, human rights abuse and poor economic conditions are very determined and seem to disregard the risks (Hatton, 2016).

Hatton's research **strategy** is based on the analysis of a database of first instance asylum applications to nineteen OECD destinations from 48 origin countries between 1997 and 2012. For each origin country, indices of civil war combat deaths (UCPD, 2015), civil liberties and political rights (Freedom House, 2015) measure war, terror (Political Terror Scale, 2015) and oppression. Real GDP per capita (Penn World Tables, 2015), and distance between origin and destination countries are also considered. A destination country's attractiveness is determined by real GDP per capita and unemployment rate, but the toughness of asylum policies is taken into account as well and measured considering the laws and regulations that could reduce asylum flows, on the basis of three components: access to territory, processing of asylum claims and welfare conditions (Hatton and Moloney, 2015).

The **econometric analysis** highlights the importance of terror in the origin country and of the lack of civil liberties. On the other hand, the effect of the lack of political rights is not significant, mainly because it prevents people from fleeing. The findings are in line with the **hypothesis** that poor economic conditions encourage asylum migration. However, this

aspect is difficult to counteract, as a 10% increase in origin country's GDP per capita causes only a 5% reduction in asylum applications. Past migration flows and distance also have a substantial impact. Applications are determined by friends and relatives, who sought asylum in the past and the volume of applications is reduced by distance, which increases the costs and the risks of irregular migration. The destination country's attractiveness is mostly determined by its unemployment rate rather than destination GDP per capita and the toughness of the asylum policies are found to be highly deterrent. Hard access and process have negative coefficients, whereas welfare conditions are positively correlated with asylum flows.

The findings highlight that asylum seekers are driven mostly by political terror and human rights abuse. Once we acknowledge that, however, finding a **solution** is not easy. Improving economic conditions would not lead to a significant change. Therefore, humanitarian organizations have put their focus on providing those trapped in refugee camps with better living conditions, especially in countries that do not have high refugee capacity. The effects of supporting and rehabilitating refugees would be greater than those from developmental aid.

The author suggests that the destination countries could improve the conditions faced by the asylum seekers during the process and implement policies which would increase their integration. Hatton believes that the **EU** should focus on three areas: firstly, existing measures and border controls need to be strengthened, in order to reduce people smuggling. Draconian measures would drastically decrease illegal immigration, prevent people from drowning in dangerous attempts of crossing the sea and gain public trust in the asylum system. Secondly, a program of resettlements of refugees from camps close to sources of violence should be developed by the Union. This would allow those most in need, the genuine refugees, to be helped, rather than those with the resources to cross the sea. Lastly, Hatton suggests the need for cooperation across developed countries to host refugees (he frames this with the economic model of public goods being underprovided). This would allow them to help more asylum seekers, by distributing them across countries, and share the social and economic costs. Such policies, together with preventing illegal immigration, would maintain the public support.

These are the practical implications drawn from the research, however, the author is well aware that these policies are only able to address a small part of a much larger problem.

Literary review

To have a better understanding of Hatton's work, we reviewed some literature answering the same research question concerning the determinants of migration.

The first article we considered is *Asylum Destination Choice: What Makes Some*

Western European Countries more Attractive than Others? (Neumayer, 2004). Neumayer uses data published by the United Nations High Commissioner for 16 Refugees (UNHCR, 2001) covering a period from 1980 to 1999 and his findings are in line with Hatton's results. This research focuses on European countries and the analysis concludes that economic attractiveness comes from income per capita, historical colonial links, common language and geographical proximity.

Bauer and Zimmermann also look at international migration in *Modelling International Migration: Economic and Econometric Issues* (1995) and find similar results, concluding that better economic conditions are the main reason driving migration. This analysis uses aggregated cross-sectional data and time series and focuses on Europe. They showed the importance of having a network in the destination country, a variable considered by Hatton as well.

Bocker and Havinga investigated the reasons for asylum applications in multiple articles. In *Asylum Applications in the European Union: Patterns and Trends and the Effects of Policy Measures* (1998a) they analysed applications in Europe through data provided by Eurostat. Like Hatton, they found that applicants historically chose the wealthiest countries in the Union. France and the UK were mentioned and the research also pointed at the possible historical colonial links. However, migration seems to have shifted to other countries, like the Netherlands, in response to the tightening of the asylum policies in Germany and France.

Bocker and Havinga, in the same year, published another paper *Asylum Migration to the European Union: Patterns of Origin and Destination* (1998b). The data was retrieved from Eurostat and the National Bureau of Statistics and focused on Europe. Asylum seekers choose destination countries based on three main factors: existing communities of compatriots, colonial bonds and the knowledge of the language. They find that the 'chain' migration, carried out thanks to the help of friends and relatives, is consistent and they highlight the effect of the historical colonial links, not really considered by Hatton (2016). Asylum seekers prefer countries with the guarantee of human rights, safety and economic wealth, especially if they are democratic and tolerant. The reception of asylum seekers and asylum policy in general does not seem to be particularly relevant.

Lastly, we considered the article '*What brings asylum seekers to the United Kingdom*?' written by Burnett (2001), which takes into account the UK and retrieves the data from the Office for National Statistics. Burnett highlights the role of civil wars, which were not significant in Hatton's dataset, and political upheavals.

We summarise the findings from our literary review in Table 1.

Paper	Source of Data	Countries	Key determinants for migration
Neumayer (2004)	UNHCR	Europe	GDP per capita, historical colonial links, common language, and geographical proximity
Bauer and Zimmermann (1995)	National Offices of Statistics	Europe	Better economic conditions
Bocker and Havinga (1998a)	Eurostat	EU	Wealthiest countries in the Union, historical colonial links, asylum policies deterrent
Bocker and Havinga (1998b)	Eurostat and the National Offices of Statistics	EU	Existing communities of compatriots, colonial bonds and the knowledge of the language
Burnett (2001)	Office for National Statistics	United Kingdom	Civil wars and political upheavals

Table 1: Literary review

Descriptive statistics

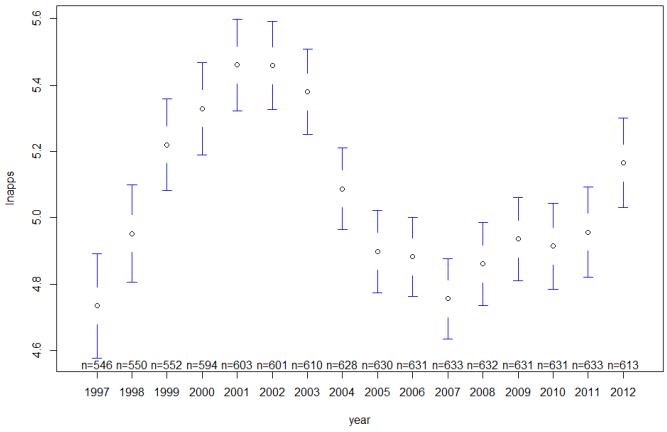
In this paragraph, we offer an overview of the dataset and see how descriptive statistics can provide a first answer to the research question. The **dataset** of unbalanced panel data used by Hatton is based on public databases. The number of applications per capita is our **dependent** variable and measures the number of people fleeing the country, in proportion to its population. The variable is expressed in logarithmic scale, in order to highlight the percentage change. The three **economic** variables are unemployment and GDP per capita in destination countries and real GDP per capita in source countries. These last two are also expressed as logarithms. Lastly, Hatton uses some indices to measure relevant **political** aspects: the Freedom House index of civil liberties and political rights (ranging from 1 to 7), the political terror scale (ranging from 1 to 5) and the Uppsala index of battle deaths in civil war. We consider 19 destination countries and 49 origin countries, listed as ordinal variables, between 1997 and 2012. The years are also expressed as ordinal variables.

The summary of our descriptive statistics can be found in Table 2.

	MEAN 97-01	MEAN 02-06	MEAN 07-12	STANDARD DEVIATION OVER WHOLE DATABASE	MIN	MAX
LOG OF ASYLUM APPLICATIONS PER CAPITA OF SOURCE POPULATION	5.1489	5.1368	4.9305	1.676	0	10.463
POLITICAL TERROR SCALE	3.5373	3.4967	3.5079	0.926	1	5
CIVIL LIBERTIES (FREEDOM HOUSE INDEX)	5.0221	4.6747	4.6198	1.317	2	7
POLITICAL RIGHTS (FREEDOM HOUSE INDEX)	5.0346	4.9564	4.9290	1.611	1	7
CIVIL WAR BATTLE DEATHS (000S)	1.0408	0.3033	0.6468	2.196	0	25.05
LOG ORIGIN COUNTRY REAL GDP PER CAPITA	7.6515	7.822	8.0503	0.961	5.192	10.233
LOG DESTINATION COUNTRY GDP PER CAPITA	10.1278	10.3497	10.5611	0.308	9.09	11.10
UNEMPLOYMENT RATE AT DESTINATION	6.8766	6.7042	7.2131	3.162	2.5	25.1
ASYLUM POLICY INDEX OVERALL	0.4859	2.8789	3.6776	3.084	-4	11
POLICY ON ACCESS	0.1599	0.9875	1.3753	0.938	-1	3
POLICY ON PROCESSING	0.0724	1.0536	1.4028	1.694	-3	6
POLICY ON WELFARE	0.2536	0.8378	0.8995	1.364	-3	6

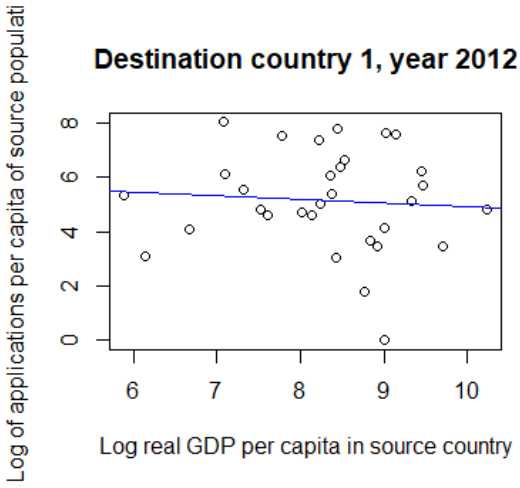
Table 2: Descriptive statistics

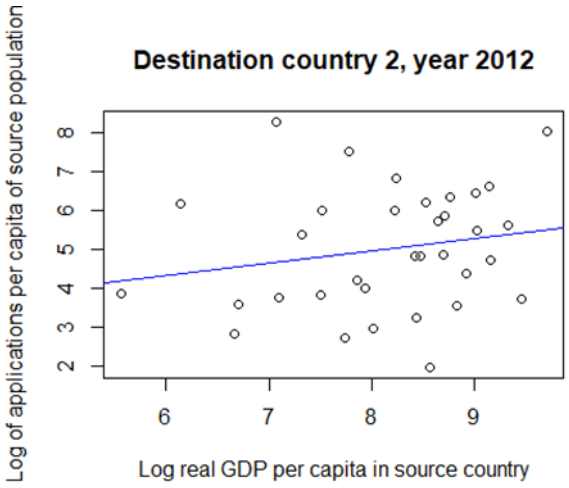
When analysing the **variations** of the variables across the periods taken into consideration, we first have to consider our dependent variable, lnapps. The logarithm of applications per capita of the source population **varies** relatively little over the years, with 170,000 decrease in the number of applications. When plotting the 95% **confidence** interval for the Log of application per capita of source population, we can see this trend of a decrease followed by an increase (See Graph 1).



Graph 1: Logarithm of applications per capita of source country, confidence interval

Then, we calculated the confidence interval over each time period and we found that the averages always lie within the bounds of the CI (See Appendix, Graph 1). If we look at the variables which could explain it, we can link it to the unemployment rate in the destination country, which after a slight decrease from 1997 to 2006, increased in the last time period. Moreover, the GDP of the source country increased as well and the index on civil war deaths decreased, improving the situation in the source country and perhaps preventing some people from choosing to seek asylum elsewhere. We plotted the **correlation** between **source GDP** and applications, and we obtained ambiguous results. For instance, in the same year, 2012, we can draw a positive correlation between the real GDP per capita in the source country and the applications per capita of source population in Country 2 and a slightly negative one for Country 1 (See Graph 2 and 3).

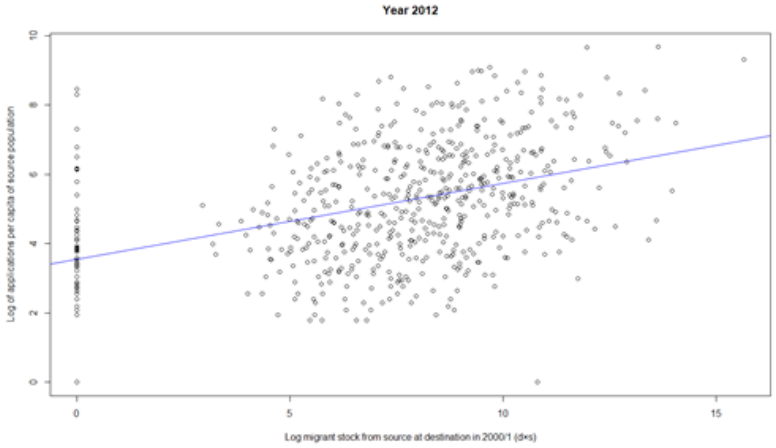




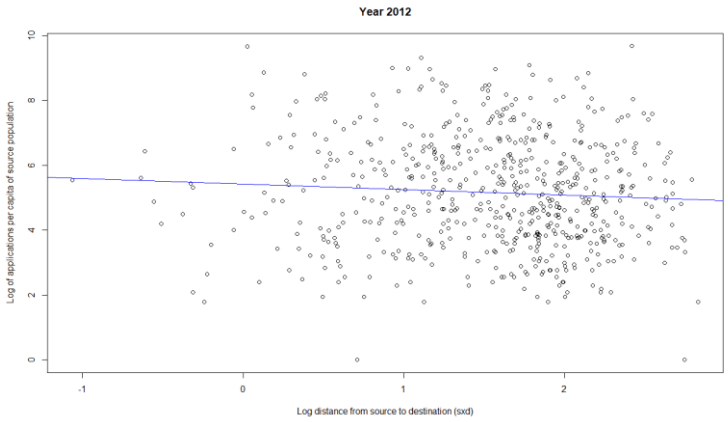
This is due to the fact that sometimes poverty could prevent people from fleeing their country. Lastly, people may have decided to stay because of the worsening of the asylum policies in the destination countries. In fact, the asylum policy index overall increased significantly from 1997 to 2012, just like all the variables related to it (policy on access, policy on processing and policy on welfare).

We assessed the **key determinants** by plotting their correlation to the logarithm of applications for two years, one at the beginning and one at the ending of the covered period. The most important one is the migrant stock from source at destination (See Graph 4), so the network that the asylum seekers can find once they reach the

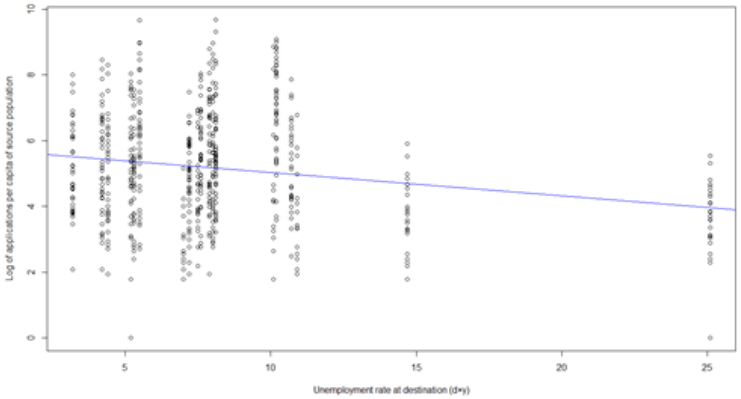
destination country: in the plot we see a strong positive relationship. We observe other strong positive correlations with distance (See Graph 5) and unemployment in the destination countries (See Graph 6). As we will see, these are the main determinants in the regression. The correlations for 2012 is here, the ones for 2000 in Appendix (Graphs 2, 3, 4)



Graph 4: Correlation between applications and stock of migrants

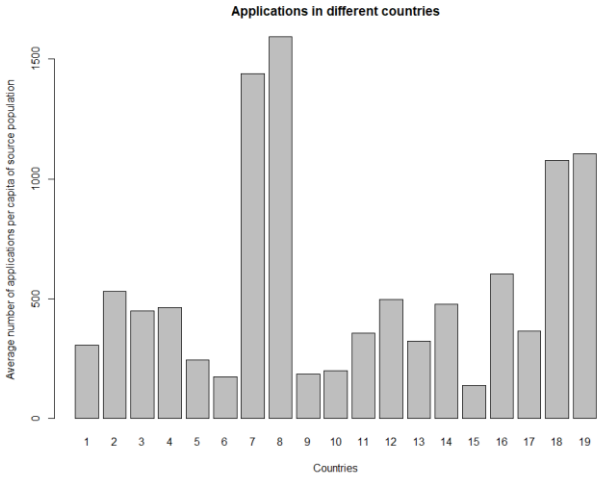


Graph 5: Correlation between applications and distance



Graph 6: Correlation between applications and stock of unemployment at destination

When it comes to variations, for checking **country wise** the **differences** in applications we used a histogram (Graph 7), in which we plotted on the x-axis the destination countries and on the vertical axis the average number of applications over the years. There are countries that received a much larger number of applications, like countries 7 and 8. It would be interesting to exclude these outliers from the analysis, to check whether the results remain consistent.



Graph 7: Countrywise difference in applications

Replication - OLS

In this paragraph, we compute the coefficients of four models using the OLS regression and we interpret their signs and significance. The four models are:

(1) $\ln apps_i = \beta_0 + \beta_1 pt_i + \beta_2 bdeaths_i + \beta_3 fhcl_i + \beta_4 fhpr_i + \beta_9 lndist_i + \beta_5 lngdpdest_i + \beta_6 lngdpsource_i + \beta_7 lnsttot_i + \beta_8 unp_i + u_i$

(2) $\ln apps_i = \beta_0 + \beta_1 pt_i + \beta_2 bdeaths_i + \beta_3 fhcl_i + \beta_4 fhpr_i + \beta_5 lngdpdest_i + \beta_6 lngdpsource_i + \beta_7 lnsttot_i + \beta_8 unp_i + \beta_9 poltot_i + u_i$

(3) $\ln apps_i = \beta_0 + \beta_1 pt_i + \beta_2 bdeaths_i + \beta_3 fhcl_i + \beta_4 fhpr_i + \beta_5 lngdpdest_i + \beta_6 lngdpsource_i + \beta_7 lnsttot_i + \beta_8 unp_i + \beta_9 polacc_i + \beta_{10} polpro_i + \beta_{11} polwel_i + u_i$

(4) $\ln apps_i = \beta_0 + \beta_1 pt_i + \beta_2 bdeaths_i + \beta_3 fhcl_i + \beta_4 fhpr_i + \beta_{12} lndist_i + \beta_5 lngdpdest_i + \beta_6 lngdpsource_i + \beta_7 lnsttot_i + \beta_8 unp_i + \beta_9 polacc_i + \beta_{10} polpro_i + \beta_{11} polwel_i + u_i$

We get the same OLS results as those provided.

	(1)	(2)	(3)	(4)
POLITICAL TERROR SCALE	0.2662*** (13.379)	0.2077*** (10.442)	0.2084*** (10.530)	0.2750*** (13.925)
CIVIL WAR BATTLE DEATHS	0.0079 (1.069)	0.0095 (1.261)	0.0087 (1.173)	0.0054 (0.747)
CIVIL LIBERTIES (FREEDOM HOUSE INDEX)	0.2239*** (9.037)	0.2160*** (8.591)	0.2056*** (8.199)	0.2082*** (8.456)
POLITICAL RIGHTS (FREEDOM HOUSE INDEX)	-0.1265*** (-6.903)	-0.1233*** (-6.635)	-0.1197*** (-6.465)	-0.1187*** (-6.529)
LOG DISTANCE FROM ORIGIN TO DESTINATION	-0.4535*** (-18.424)			-0.4672*** (-18.928)
LOG DESTINATION COUNTRY GDP PER CAPITA	0.0804 (1.418)	0.0312 (0.512)	0.0813 (1.301)	0.3393*** (5.397)
LOG ORIGIN COUNTRY GDP PER CAPITA	-0.1486*** (-7.972)	-0.0309. (-1.727)	-0.0346* (-1.943)	-0.1479*** (-8.003)
LOG MIGRANT STOCK IN 2000/1 FROM ORIGIN AT DESTINATION	0.2359*** (47.551)	0.2210*** (43.943)	0.2241*** (44.619)	0.2331*** (47.056)
UNEMPLOYMENT RATE AT DESTINATION	-0.0268*** (-5.197)	-0.0347*** (-6.627)	-0.0318*** (-5.478)	-0.0333*** (-5.844)
ASYLUM POLICY INDEX OVERALL		-0.0421*** (-7.584)		
POLICY ON ACCESS			-0.0964*** (-5.119)	-0.0694*** (-3.741)
POLICY ON PROCESSING			-0.1082*** (-9.378)	-0.1236*** (-10.881)
POLICY ON WELFARE			0.0701*** (5.400)	0.0500*** (3.908)

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
In brackets the T statistics

Table 3: OLS estimation results

For the interpretation of the coefficients, many are statistically significant at level p<0.001. Before arguing that these coefficients are important determinants of asylum applications, we note that the standard errors are extremely low, and thus the **t-statistics** give extremely high results. For example, the t-value for the variable Log migrant stock from source at destination in 2000/1 returns values above 43 in all four models. We are in a panel data context, and this can be a signal that conditions of **homoskedasticity** and lack of **correlation** between X_i and errors might not hold. Therefore, the estimators may still be unbiased, however, heteroskedasticity and autocorrelation make the OLS estimators no longer the most efficient ones (BLUE). Another signal is that the coefficient for the Freedom House Index on political rights is statistically significant but has a counterintuitive sign, revealing a problem with the econometric method. In fact, a studentized Breusch-Pagan test rejects the null hypothesis of homoskedasticity for all the four models, and a Durbin-Watson test rejects the null hypothesis of lack of autocorrelation for all the models. Moreover, the dependent variable is in logarithmic form, so it can only be non-negative. The assumption of normality $u \sim N(0, \sigma^2)$ does not hold, which leads us to believe that the t-statistics are not reliable.

We **interpret** then only the variables statistically significant with t-tests based on **HAC** standard errors (See Appendix, Table 1): these alter the standard errors, leaving the coefficients unchanged. T-statistics computed with HAC are only asymptotically valid, but here we have a large sample. The discouraging effect of the distance between origin and destination country should be taken into consideration by policy makers. Its high coefficient should be interpreted as a proportion between the percentage increase of the number of applications and the percentage increase of the distance, since both the dependent and the independent variables are in logarithmic form. As Hatton (2013, p.442) points out, Political Terror is extremely relevant also in our OLS models. Another relevant **origin country** effect is the lack of civil rights: ceteris paribus, an increase in lack of civil rights on a scale of 1 to 7 pushes up applications by around 20 percent. Interestingly, the reasoning does not apply to political rights: apart from not being significant using HAC standard errors, we get an unexpectedly negative coefficient, so that the lack of political rights decreases asylum applications. Perhaps, the effects of the three subcategories included in the index (Electoral

Process, Political Pluralism and Participation, Functioning of Government) have contrasting effects and disentangling them can be a useful strategy to better understand the effects of the single subcategory.

Regarding the **destination country** effects, the two HAC statistically significant variables are the policy on asylum processing and the number of people from the origin country already living in the destination country. The index for asylum processing is based on five subcategories (Hatton & Moloney, 2015): definition of a refugee, humanitarian category, manifestly unfounded claims, expedited procedures (fast and cost-effective resolution of disputes), scope for appeals. If a country gets stricter rules in these fields, the index will scale up, thus the coefficient will be negative. The definition of strictness is based on new laws, but its valuation is intrinsically subjective.

Clearly, no model uses both the overall policy index and the three specific policy indices, as it would incur in a perfect **collinearity** problem. The indices have an exact linear relationship $poltot = polpro + polacc + polwel,$ thus it would be impossible to untie the effects of the overall index and of the specific ones. As we will argue in the discussion, migrant stock is one of the most important determinants in our model. It is not surprising that people prefer to move to a place in which they can find a community of their nationality, as these people can understand their language and culture and potentially help them to find a job (Bocker and Havinga, 1998).

Finally, in terms of sensitivity analysis, it is interesting to notice that in the OLS model the origin GDP per capita is significant at $p<0.01$ level only when distance is included in the model, otherwise it is just significant at a $p<0.1$ level. In this case, the omitted variable of the distance arguably makes the destination GDP endogenous and **biased** towards zero: the biased coefficient is lower and thus no longer statistically significant.

These were our remarks concerning the OLS estimation results. We can then proceed in our analysis with a different model.

Replication – Fixed Effects

In this paragraph, we will analyse the FE model and comment on the key determinants. The estimation model used by Hatton (2013) is the **fixed effect** model (unobserved variables which do not change over time are fixed), and thus allows all the other explanatory variables to be correlated with unobservable ones that are constant over time (α_i). When we use origin country as a fixed effect, we account for all factors affecting the number of asylum applications which are roughly constant over time but change in different origin countries, like demographic features of the population and geographical position. The same goes for origin × destination: in this case, we can account for roughly constant specific relationships between an origin and a destination country, as the distance, the number of origin country's

citizens living in the destination country, but also historical factors like the existence of a colonial link. Conversely, origin × year accounts for events that happened in an origin country in a specific year: for example, a war or the rise of a dictatorship. These are the four **models**:

(1) $\ln apps_i = \beta_0 + \beta_1 pt_i + \beta_2 bdeaths_i + \beta_3 fhcl_i + \beta_4 fhpr_i + \beta_9 lndist_i + \beta_5 lngdpdest_i + \beta_6 lngdpsource_i + \beta_7 lnsttot_i + \beta_8 unp_i + \beta_9 year + \beta_{10} dcode + \beta_{11} scode + u_i$

(2) $\ln apps_i = \beta_0 + \beta_1 pt_i + \beta_2 bdeaths_i + \beta_3 fhcl_i + \beta_4 fhpr_i + \beta_5 lngdpdest_i + \beta_6 lngdpsource_i + \beta_7 lnsttot_i + \beta_8 unp_i + \beta_9 poltot_i + \beta_{10} year + \beta_{11} dscode + u_i$

(3) $\ln apps_i = \beta_0 + \beta_1 pt_i + \beta_2 bdeaths_i + \beta_3 fhcl_i + \beta_4 fhpr_i + \beta_5 lngdpdest_i + \beta_6 lngdpsource_i + \beta_7 lnsttot_i + \beta_8 unp_i + \beta_9 polacc_i + \beta_{10} polpro_i + \beta_{11} polwel_i + \beta_{12} year + \beta_{13} dscode + u_i$

(4) $\ln apps_i = \beta_0 + \beta_1 pt_i + \beta_2 bdeaths_i + \beta_3 fhcl_i + \beta_4 fhpr_i + \beta_{12} lndist + \beta_5 lngdpdest_i + \beta_6 lngdpsource_i + \beta_7 lnsttot_i + \beta_8 unp_i + \beta_9 polacc_i + \beta_{10} polpro_i + \beta_{11} polwel_i + \beta_{12} dcode + \beta_{13} scodeyear + u_i$

	(1)	(2)	(3)	(4)
POLITICAL TERROR SCALE	0.2142***	0.2210***	0.2207***	
CIVIL LIBERTIES (FREEDOM HOUSE INDEX)	0.2846***	0.2895***	0.2916***	
POLITICAL RIGHTS (FREEDOM HOUSE INDEX)	-0.0436	-0.0503	-0.0492	
CIVIL WAR BATTLE DEATHS	0.0123	0.0103	0.0106	
LOG ORIGIN COUNTRY GDP PER CAPITA	-0.5172*	-0.5334**	-0.5421**	
LOG MIGRANT STOCK IN 2000/1 FROM ORIGIN AT DESTINATION	0.2256***			0.2262***
LOG DISTANCE FROM ORIGIN TO DESTINATION	-0.7773***			-0.7678***
LOG DESTINATION COUNTRY GDP PER CAPITA	0.1781	0.0660	-0.1223	-0.0428
UNEMPLOYMENT RATE AT DESTINATION	-0.0246*	-0.0238*	-0.0240*	-0.0286***
ASYLUM POLICY INDEX OVERALL		-0.0436***		
POLICY ON ACCESS			-0.1145**	-0.1102***
POLICY ON PROCESSING			-0.0997***	-0.1032***
POLICY ON WELFARE			0.0490 .	0.0341

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 3: FE estimation results

We got the same results as Hatton. However, to get these results we had to include all the dummy variables, without excluding the first one to avoid problems of multicollinearity. If we do not put the first dummy in the model, we get slightly different results. We should carefully consider if there is a problem of multicollinearity, as the lack of it is one of the conditions to run a regression.

We allow the intercept to differ across the years by using year and destination code **dummies**, to reflect the different distributions of population in different time periods. We did not report the coefficients of the dummies, even if some of them are statistically significant.

Hatton **comments** on the variables that have the biggest effect on the number of applications. Although the political terror index is not the most relevant among the significant variables, Hatton considers it rather important and chooses to discuss extensively this determinant. The distance from origin to destination has the largest coefficient among the statistically significant values. However, clearly, this determinant cannot be manipulated by policymakers. Variables related to the origin country usually have a bigger impact than those related to the destination country, suggesting that the **push** effect can be more important than the pull one. The destination country GDP is not significant, while the origin country GDP has a negative effect, significant at the 5% level. The fact that mainly relatively rich people have enough money to engage in a long and unsure trip is widely recognized (EC, 2015), but arguably origin GDP is not an effective variable to capture it.

Still on origin effects, it is not easy to capture through quantitative variables the political upheavals in the origin countries, but Hatton (2013) succeeds using the political terror scale and civil liberties (Freedom House index), both significant and both with a

considerable effect on the dependent variable. Instead, civil war battle deaths and the Freedom House index of political rights seem to be irrelevant. As Hatton rightfully points out, paradoxically political rights have an ambiguous effect on the number of asylum seekers: the incentive to flee is high, but the ability to do so is not. Since the effect is close to zero, the sign can also be wrong due to error sampling (Wooldridge, 2018, p.655). The same reasoning can be applied to the number of civil war deaths: people might want to leave the country, but something could prevent them from doing so (for example, being conscripted and knowing that deserting leads to a death sentence).

Economic destination country effects do not seem to be fundamental: the log destination country GDP per capita is not statistically significant, while the unemployment rate is, but the magnitude of the coefficient is considerably small. Statistical significance can differ from economic significance, since, if the size of beta is small, the effect of x on y is negligible, we should be careful in attributing much importance to that variable (Wooldridge, 2018, p.655). By contrast, the **destination** country effects on a **social** level seem rather important: destinations' policies on access and on processing are probably known to migrants. We can assume that there is an exchange information, at least on an informal level, among migrants on the same journey. The refugees are therefore to a certain extern aware of the process they are going to face to receive the refugee status. This knowledge has an impact on their decision to move to a certain country. Overall, however, the most important effect concerning the destination choice is the network of compatriots living there.

This was our commentary regarding the Fixed Effects estimation results. We can then continue with a sensitivity analysis to test the results we got with this estimation model.

Sensitivity Analysis

In this paragraph, we will check the robustness of estimated effects and significance. Firstly, we dropped the two **outlier** countries 7 and 8, which have a number of applications much higher than the other destination countries (cfr. Graph 7). For all the four columns, there are differences only at the second or third digit after the decimal point in the coefficients, and the effect on statistical significance is minimal (See Appendix, Table 2). The only difference is that the index for policies on welfare is statistically significant at $p>0.05$ instead of $p>0.1$ in the last two models. Enderlein (1987) considers outliers values that deviate so much from other observations that one might suppose a different underlying sampling mechanism. In fact, using only the two outliers, we get completely different results with regards to coefficients and statistical significance. (See Appendix, Table 3)

The models, both OLS and FE, tested with Breusch-Pagan test, are **heteroskedastic**. They also result to be autocorrelated from the Durbin-Watson tests. In this test, the null

hypothesis is that the residuals are not autocorrelated and it is rejected. However, Hatton used standard errors that are clustered by origin country in the first model, by origin × destination in the second and third one, by origin × year in the fourth one. Clustered standard errors allow for forms of serial correlation and heteroskedasticity in panel data: they require a large cross section (N) and not too large time series (T), and we have both here (N=10128, T=16), therefore, we can consider these two issues solved.

We can speculate that the variable "log destination country GDP per capita" is **endogenous**. It may be reasonable that other factors that have an impact on the number of asylum seekers in a country are omitted and positively correlated with the destination GDP. We can think of other factors in a destination country that can contribute to its attractiveness but are not taken into account, for instance the quality of the health system, which can be estimated with the percentage of government spending for health. To disentangle this factor from the error term, we would need an additional dataset. This, however, implies that the GDP could be biased upward due to this omitted variable. Overall government spending could also be considered an omitted variable, as it is positively correlated with GDP per capita and negatively related to unemployment rate. Therefore, the direction of the bias is ambiguous. If we were to use an **IV** to solve this problem of endogeneity, we would have to find a variable, which is correlated with the destination GDP but not with the error term. A good candidate would be net exports. This has a direct impact on the GDP and we can realistically assume it has no effect on the number of asylum seekers, as they probably do not know a destination country's imports and exports. Since we do not have this data, without any IV we cannot **test for exogeneity**, thus we cannot know if one of the main assumptions in this model holds.

When we use an imprecise measure of an economic variable in a regression model, our model contains **measurement error**. For instance, we would like to know exactly which factor of the political and social situation in the origin country makes people migrate, but we are obliged to use some indices, because otherwise we would have to account for too many different factors. Consequently, we cannot estimate the direction of this bias.

After having checked the robustness of our results, we can discuss the findings.

Discussion

In this paragraph, we will identify strengths and weaknesses of Hatton's analysis and discuss them.

Causal relationships are established with potential sources of **bias**, because of numerous reasons. Firstly, we cannot be sure that asylum seekers intended to be in the destination countries in which they found themselves. Therefore, when we say that for instance the real GDP per capita of the destination countries is not statistically significant, it could be due to this bias. Perhaps people from Middle East would like

to go to a country with a higher GDP, like the United States, but their choice is limited and they actually obtain the refugee status in a country with a lower GDP per capita, for instance Lebanon. To assess this, we would need an ideal empirical strategy that would test intentions. However, this would be difficult to carry out, if not infeasible. Hatton does not consider this distinction and the **reader** should acknowledge this difference. Further, Hatton chose to only include those dyads of origin and destination countries that had a number of applications exceeding 300 over 16 years. Consequently, he excluded the lower bound outliers but not the upper bound ones (for instance, destination countries 7 and 8), altering the distribution.

Comparing OLS and FE models, we notice that problems in the OLS model are mainly about efficiency and inference rather than unbiasedness, so the coefficients of OLS and FE are different, but usually this difference appears from the second decimal, while the direction of the effect is unaffected. As we pointed out, both OLS and FE models are heteroskedastic and with problems of serial correlation: in the former, we computed HAC, while in the latter Hatton (2016) clusters standard errors by origin country.

In both models the variables are overall significant, in the model with fixed effects, the dependent variables are overall significant both for the full model and for the projected model for all the columns. We can then say that the **independent** variables are **appropriate**, even if some problems arise. First, Hatton (2016) never refers to the base year of the real GDPs. Moreover, he considers the origin country GDP, but he did not consider in which income range the asylum seekers were. If these migrants were in the wealthiest top 5% of the population, the origin country GDP does not reflect their real condition. There are alternative indices to GDP per capita, to compute well-being, for example the Thriving Places Index (Brunner, 2017). They can perhaps better estimate the pull effect, as they account also for how the wealth is distributed, rather than using the per capita variable, thus only dividing for the number of citizens. Moreover, the analysis would be more precise if we had the migrant stock for every year, not just for 2000/01, but this data cannot be easily retrieved. Finally, he does not justify why he chose those indices. The choice of some of them is arguable: for example, the Freedom House NGO can be accused of favouritism towards the US, as one of its goals is “Promoting U.S. Leadership” (Freedom House, 2015). We can also remark that the **dependent** variable used is **appropriate**, since using the logarithmic form, he is able to measure the percentage increase in the number of applications. This variable is expressed as a ratio of number of applications to population of the source country. Therefore, lnapps accounts for the differences in population density and provides an informative value, which does not grow disproportionally for populous nations.
In terms of data, Hatton (2016) succeeded in finding data from all over the world, but without knowing the relation between

countries and country codes, it is difficult to really interpret the results. For example, some States are linked through a colonial history or share the same language, but we are not able to identify them. We can overcome this obstacle by using a dummy variable to identify former colonial links, however Hatton tackles this issue by fixing the effects origin × destination in the 2nd and 3rd models.

We can try to **improve** the **analysis** using the **current dataset**. For example, we said that the GDP of the origin country has an ambiguous effect on the number of applications. People from very poor countries are not able to migrate because of the cost of the process. On the other hand, people from rich countries do not have the need to seek asylum. It seems reasonable to assume that the origin country GDP can have a **parabolic** form. To test it, we include the variable $lngdpdest^2$ in the models with the fixed effects, but this variable is not statistically significant (See Appendix, Table 4). However, including $lngdpdest^2$, also the logarithm of the origin GDP is no longer significant.

We can test exclusion for variables which are highly correlated with a heteroskedasticity-robust version of the **joint significance** test. However, we can reject the null hypothesis that both coefficients are zero at any level of significance for political terror and civil war battle deaths (in the first three models), and for the number of people already in the country and the distance (in the first and fourth model). When it comes to the relationship between the policy index on welfare and the destination GDP, we reject the null hypothesis of joint non-significance at least at level p>0.01 for the model with fixed effects destination × origin and the one with the fixed effects are origin × year.

We do not have problems of seasonality, since the time periods are years. However, the **trend** variable we introduced is statistically significant at level p>0.01 in the second model, and in the third one at any level used in practice, while it is not for the first and the fourth ones (See Appendix, Table 5). Thus, in some cases, it would be plausible to de-trend the data. Reasonably, some of the explanatory variables, for example the GDPs, have a positive trend. It could be reasonable to allow for a **finite distributed lag**, since for example an increase in political terror in a year does not have an effect on migration in the same year but in the following ones. Nonetheless, the variables related to origin country (pt, lngdpsource, fhcl, fhpr, bdeaths) lagged by one year and two years are not statistically significant, except political terror of two years before (See Appendix, table 6 and 7)

If we analyse the differences between our replication and Hatton's, we used a more precise significance level, while he used only the two standard significance levels of 5% and 10%.

This terminates our discussion, finally we would like to suggest some topics for future research in this field.

Extension

In terms of **further research,** it would be interesting to include data from economic migrants, to see which variables would still be significant in this case. We would expect origin and destination GDP and especially unemployment to be the most relevant variables. The dataset used by Hatton could have perhaps been more representative, if it were not limited to first instance applications, but if it were to consider those who had tried and been rejected in the past. Other interesting determinants to include could be the ease of finding employment, after that one has obtained the refugee status. This could be measured, for example, by the number of months of unemployment. We could also consider the level of integration of one's ethnical community in the destination country. For instance, this could be reflected by the number of people of that community working and being involved in the public life (i.e. volunteering or being part of an association).

We need to keep in mind that our suggestions could be to some extent helpful to design better policies addressing the issue, however, overcoming the obstacles will be a great challenge.

References:

Bauer, T. and Zimmermann, K., (1994). *Modelling International Migration: Economic and Econometric Issues*, in Rob van der Erf and Liesbeth Heering(eds) *Causes of International Migration*, pp. 95-115. Brussels: Statistical Office of the European Communities.

Böcker, A. and Havinga, T., (1998). *Asylum Applications in the European Union: Patterns and Trends and the Effects of Policy Measures*, Journal of Refugee Studies 11(3): 245-66.

Böcker, A., (1998). *Asylum Migration to the European Union: Patterns of Origin and Destination*. Luxembourg: Office for Official Publications of the European Communities

Davenport, C. A., Moore W. H., and Poe S.C. (2003). *Sometimes You Just Have to Leave: Domestic Threats and Forced Migration, 1964–1989*. International Interactions: Empirical and Theoretical Research Relations 29 (1): 27–55.

Enderlein, G. (1987). *Hawkins, Dm: Identification of Outliers*. *Chapman and Hall, London–New York 1980*, Biometrical Journal 29 (2): 198–98.

European Commission, DG Migration & Home Affairs. (2015) *A study on smuggling of migrants Characteristics, responses and cooperation with third countries.* Final Report September 2015 Retrieved from: https://ec.europa.eu/home-affairs/sites/homeaffairs/files/what-we-do/networks/european_migration_network/reports/docs/emn-studies/study_on_smuggling_of_migrants_final_report_master_091115_final_pdf.pdf

Hatton, T. J., and Moloney J. (2015). *Applications for Asylum: Modelling Asylum Claims by Origin and Destination*. CEPR Discussion Paper 10678.

Moore, W. H., and. Shellman S. M. (2007). *Whither Will They Go? A Global Study of Refugees' Destinations, 1965–1995.* International Studies Quarterly 51 (4): 811–34.

Neumayer, E., (2004). *Asylum Destination Choice: What Makes Some West European Countries More Attractive Than Others?*. European Union Politics - EUR UNION POLIT, 5, pp.155-180.

Wooldridge, J. M., (2012). *Introductory econometrics: a modern approach.* Mason, Ohio :South-Western Cengage Learning,

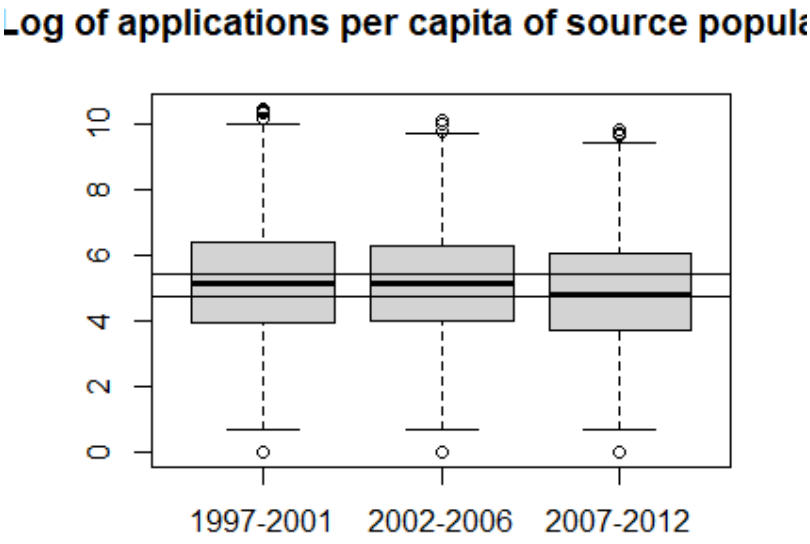
Uppsala Conflict Data Program, 2015. "Civil war combat deaths". Retrieved from: https://ucdp.uu.se/downloads/

Freedom House, 2015. "Freedom in the world Annual Report - 2015". Retrieved from: https://freedomhouse.org/sites/default/files/01152015_FIW_2015_final.pdf

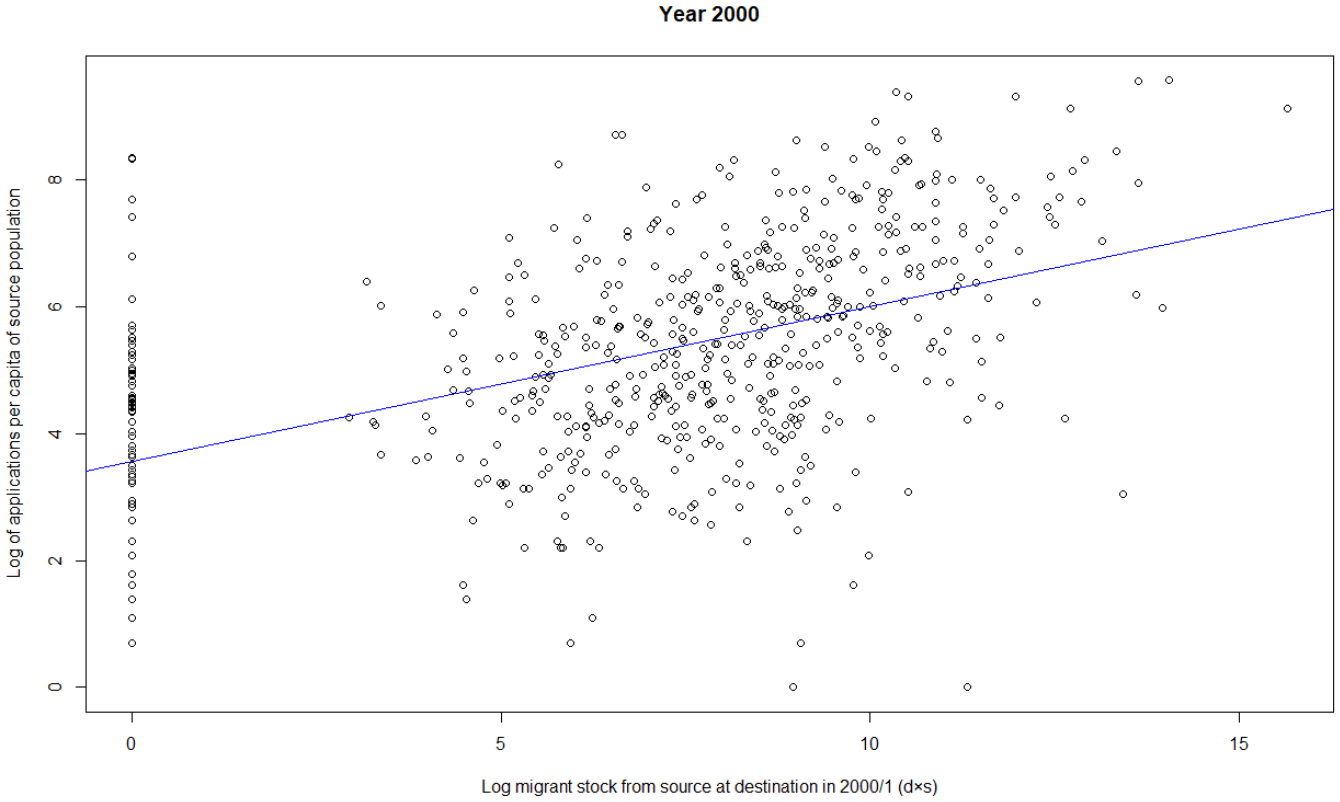
Penn World Tables, 2015. "Real GDP using national-accounts growth rates (RGDPNA) 1950-2017". Retrieved from: https://febpwt.webhosting.rug.nl/Dmn/AggregateXs/PivotShow

University of North Carolina at Asheville. "2015 PTS Release". Retrieved from: http://www.politicalterrorscale.org/Data/

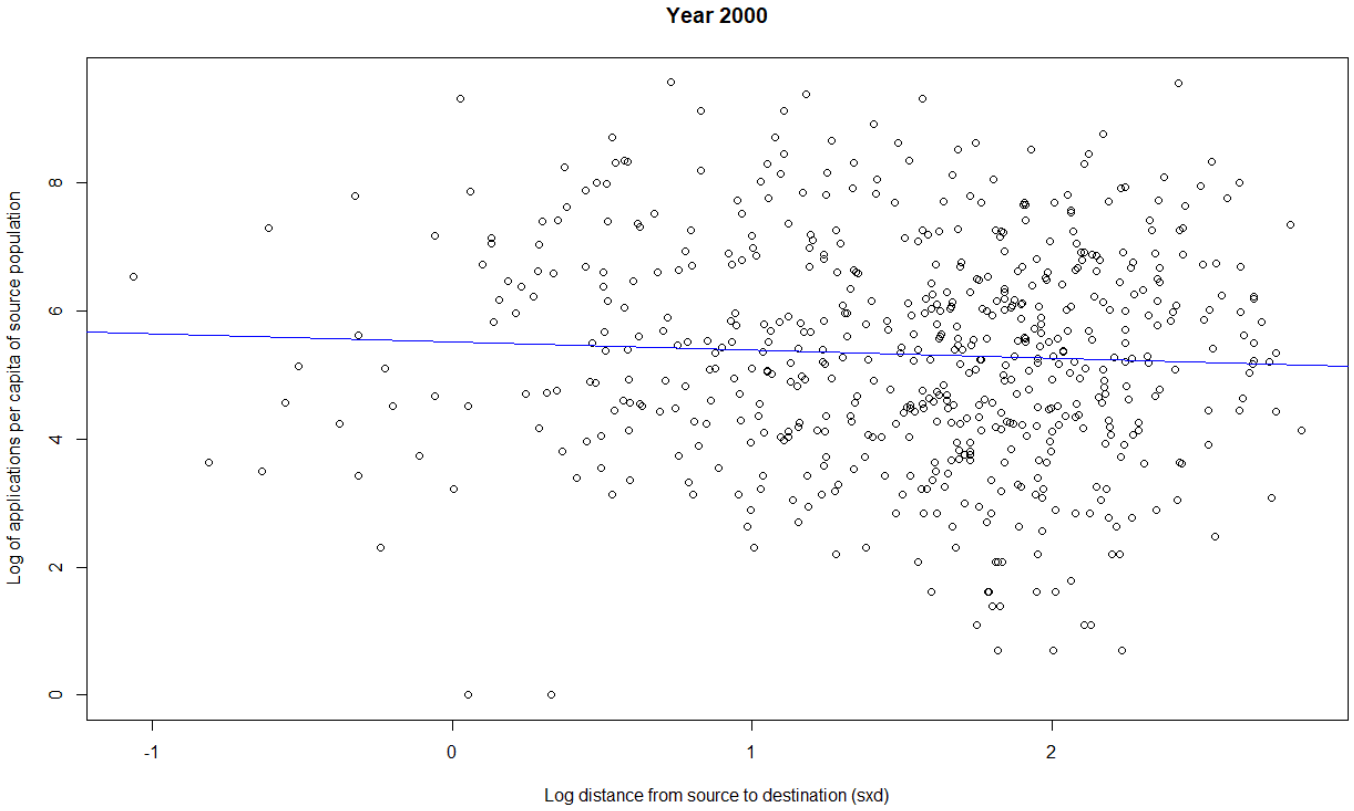
Graph 1. Confidence interval over three time periods.



Graph 2: correlation between Log of applications per capita of source population and Log migrant stock from source at destination in 2000/1 for year 2000.



Graph 3: correlation between Log of applications per capita of source population and Log distance from source to destination for year 2000.



Graph 4: correlation between Log of applications per capita of source population and unemployment in the destination country.

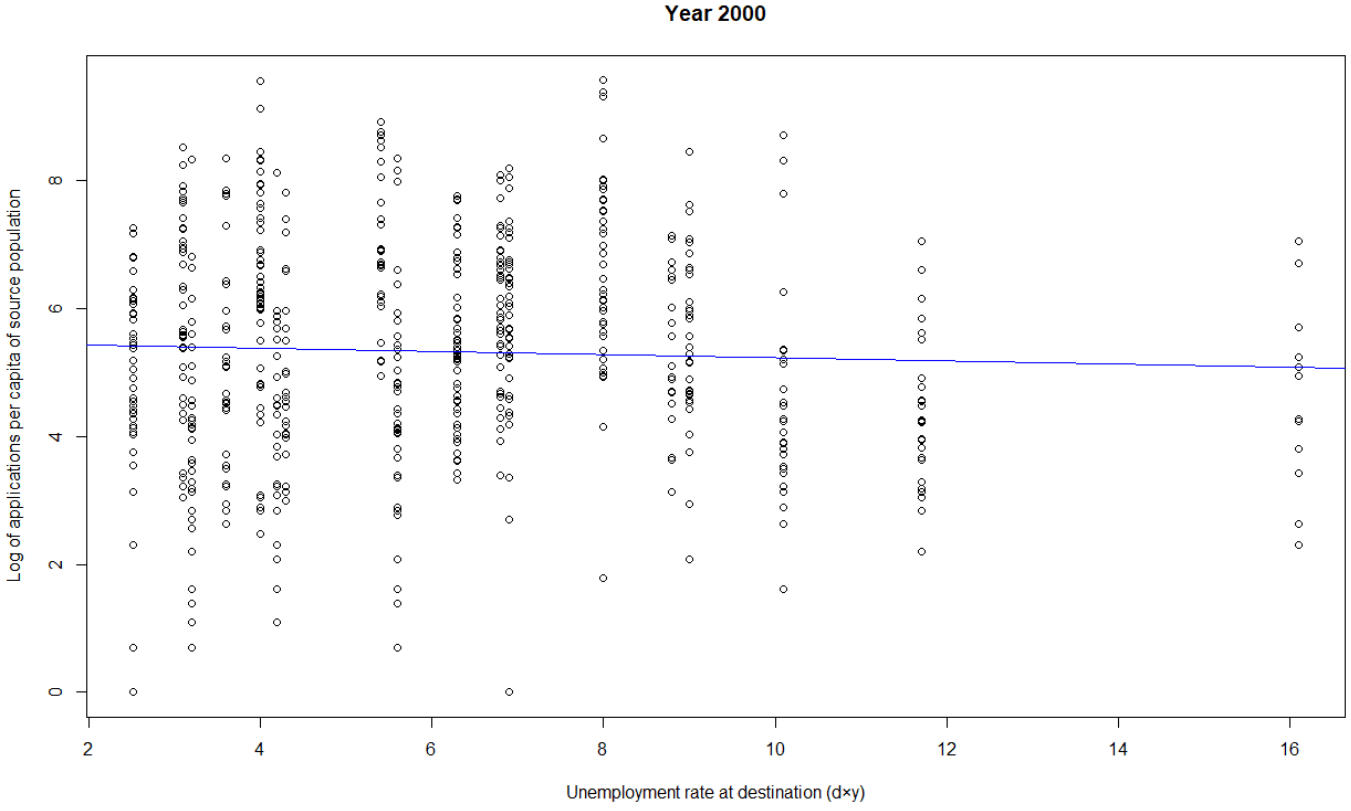


Table 1: OLS with HAC standard errors. In brackets, HAC T-statistics

Political terror scale	0.2661*** (5.6004)	0.2077*** (4.1432)	0.2084*** (4.1618)	0.2750*** (5.8062)
Civil war battle deaths	0.0079 (0.7708)	0.0095 (0.8969)	0.0087 (0.8310)	0.0054 (0.5314)
Civil liberties (Freedom House index)	0.2238*** (4.4118)	0.2160*** (4.1798)	0.2056*** (3.9777)	0.2082*** (4.1256)
Political rights (Freedom House index)	-0.1264** (-3.1187)	-0.1233*** (-3.0075)	-0.1197*** (-2.9256)	-0.1187*** (-2.9709)
log distance from origin to destination	-0.4535*** (-5.3816)			-0.4672*** (-5.5867)
log destination country GDP per capita	0.0803 (0.5898)	0.0312 (0. 2042)	0.0813 (0.5155)	0.3393*** (2.0803)
log origin country GDP per capita	-0.1485** (-2.8181)	-0.0309 (-0.5808)	-0.0346* (-0.6626)	-0.1479*** (-2.8374)
log migrant stock in 2000/1 from origin at destination	0.2358*** (13.9686)	0.2210*** (13.0687)	0.2241*** (13.4064)	0.2331*** (14.0609)
Unemployment rate at destination	-0.02677* (-2.0700)	-0.0347*** (-2.7275)	-0.0318*** (-2.5238)	-0.0333*** (-2.6400)
Asylum policy index overall		-0.0421*** (-2.8746)		
Policy on access			-0.0964*** (-2.1755)	-0.0694*** (-1.6359)
Policy on processing			-0.1082*** (-3.6421)	-0.1236*** (-4.3024)
Policy on welfare Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1			0.0701*** (1.8371)	0.0500*** (1.3401)

Table 2: FE models without outliers

Political terror scale	0.2268 ***	0.2312***	0. 2300***	
Civil liberties (Freedom House index)	0. 2797***	0.2839***	0. 2875***	
Political rights (Freedom House index)	-0.0434	-0.0492	-0.0478	
Civil war battle deaths	0.0104	0.0090	0. 0094	
log origin country GDP per capita	-0.5181*	-0.5365**	-0.5424**	-0.1827
log migrant stock in 2000/1 from origin at destination	0.2237***			0.2241***
log distance from origin to destination	-0.7989***			-0.7881***
log destination country GDP per capita	0.2620	0.1485	-0.1616	0.2262***
Unemployment rate at destination	-0.0257*	-0.0248*	-0.0281*	-0.0356***
Asylum policy index overall		-0.0592***		
Policy on access			-0.1682**	-0.1625***
Policy on processing			-0.1231***	-0.1360***
Policy on welfare Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1			0.05844 .	0.0459 .

Table 3: FE models computed only on outliers

Political terror scale	0.1359 .	0.1393 *	0.1458 *	
Civil liberties (Freedom House index)	0.3271 **	0.3255 ***	0.3216 ***	
Political rights (Freedom House index)	-0.0517	-0.0599	-0.0603	
Civil war battle deaths	0.0227	0.0169	0.0152	
log origin country GDP per capita	-0.5019	-0.5149	-0.5097	
log migrant stock in 2000/1 from origin at destination	0.2112 *			0.1990 ***
log distance from origin to destination	-0.1083			-0.1693
log destination country GDP per capita	-15.3730 ***	4.7380 **	6.6859 ***	4.7253
Unemployment rate at destination	-0.2156 ***	-0.0874 *	-0.2616 ***	-0.2422 ***
Asylum policy index overall		0.7003 ***		
Policy on access			-0.0836	-0.0597
Policy on processing			1.6524 ***	1.4743 ***
Policy on welfare Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1			0.9351 ***	0.8371 ***

Table 4: FE models with a parabolic variable

Political terror scale	0.2102 ***	0.2181 ***	0.2175 ***
Civil liberties (Freedom House index)	0.2815 ***	0.2871 ***	0.2891 ***
Political rights (Freedom House index)	-0.0384	-0.0465	-0.0449
Civil war battle deaths	0.0120	0.0101	0.0104
log origin country GDP per capita	0.5096	0.1890	0.2668
log migrant stock in 2000/1 from origin at destination	0.2256 ***		
log distance from origin to destination	-0.7773 ***		
log destination country GDP per capita	0.1732	0.0629	-0.1267
lngpdsource^2	-0.0645	-0.0454	-0.0508
Unemployment rate at destination	-0.0248 *	-0.0239 *	-0.0240 *
Asylum policy index overall		-0.0462 ***	
Policy on access			-0.1147 **
Policy on processing			-0.0998 ***
Policy on welfare Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1			0.0497 .

Table 5: FE models with a trend

Trend	-0.0006	0.0720 **	0.0851 ***	-0.0006
Political terror scale	0.2142 ***	0.2210 ***	0.2207 ***	
Civil liberties (Freedom House index)	0.2844 ***	0.2895 ***	0.2916	
Political rights (Freedom House index)	-0.0435	-0.0503	-0.0492	
Civil war battle deaths	0.0123	0.0103 ***	0.0102 ***	
log origin country GDP per capita	-0.5176 *	-0.5334 *	-0.5239 *	
log migrant stock in 2000/1 from origin at destination	-0.7687 ***			0.2254 ***
log distance from origin to destination	0.2249 ***			-0.7601 ***
log destination country GDP per capita	0.1766	0.0660	-0.1222	-0.0447
Unemployment rate at destination	-0.0246 *	-0.0238 *	-0.0239 *	-0.0286 *
Asylum policy index overall		-0.0454 ***		
Policy on access			-0.1145 **	-0.1112 ***
Policy on processing			-0.0997 ***	-0.1031 ***
Policy on welfare			0.0490 .	0.0342 .

Table 6: FE models with a lag (t-1)

Political terror scale	0.1802 ***	0.1941 ***	0.1941 ***
pt_{t-1}	0.0601 .	0.0627 .	0.0613 .
Civil war battle deaths	0.0091	0.0119	0.0122 ***
$bdeaths_{t-1}$	-0.0094	-0.0091	-0.0089
Civil liberties (Freedom House index)	0.2910 ***	0.2860 ***	0.2879 ***
$fhcl_{t-1}$	-0.0184	-0.0140	-0.0125
Political rights (Freedom House index)	-0.0351	-0.0245	-0.0366
$fhpr_{t-1}$	-0.00155	-0.0380	-0.0250
log distance from origin to destination	-0.7762 ***		
log destination country GDP per capita	0.1988	0.0806	-0.1113
log origin country GDP per capita	-0.5764 **	-0.5652*	-0.5740 *
$lngdpsource_{t-1}$	0.0682	0.0583	0.0588
log migrant stock in 2000/1 from origin at destination	0.2261 ***		
Unemployment rate at destination	-0.0245*	-0.0237 *	-0.0238 *
Asylum policy index overall		-0.0456 ***	
Policy on access			-0.1134 **
Policy on processing			-0.1013 ***
Policy on welfare Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1			0.0530 .

Table 7: FE models with a lag (t-2)

Political terror scale	0.1876 ***	0.2001 ***	0.1997 ***
pt_{t-2}	0.0613	0.0677 .	0.0668 .
Civil war battle deaths	0.0091	0.0063	0.0067 ***
$bdeaths_{t-2}$	-0.0094	-0.0135	-0.0134
Civil liberties (Freedom House index)	0.2870 ***	0.2827 ***	0.2850
$fhcl_{t-2}$	-0.0622	-0.0557	-0.0543
Political rights (Freedom House index)	-0.0408	-0.0470	-0.0456
$fhpr_{t-2}$	-0.0040	-0.0135	-0.0141
log distance from origin to destination	-0.7762 ***		
log destination country GDP per capita	0.1841	0.0668	-0.1215
log origin country GDP per capita	-0.5416 *	-0.5348*	-0.5434 *
$lngdpsource_{t-2}$	0.0521	0.0456	0.0461
log migrant stock in 2000/1 from origin at destination	0.2259 ***		
Unemployment rate at destination	-0.0250*	-0.0243*	-0.0244 *
Asylum policy index overall		-0.0458 ***	
Policy on access			-0.1122 **
Policy on processing			-0.1007 ***
Policy on welfare Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1			0.0510 .