

Migration patterns, friendship networks, and the diaspora: the potential of Facebook’s Social Connectedness Index to anticipate displacement patterns induced by Russia’s invasion of Ukraine in the European Union

Umberto Minora^{1*}, Martina Belmonte¹, Claudio Bosco¹, Drew Johnston², Eugenia Giraudy³, Stefano M Iacus⁴, and Francesco Sermi¹

¹*Joint Research Centre, European Commission*

³*Core Data Science, Meta*

²*Department of Economics, Harvard University*

⁴*Institute for Quantitative Social Science, Harvard University*

**Corresponding author: umberto.minora@ec.europa.eu*

Abstract

The conflict in Ukraine is causing large-scale displacement in Europe and in the World. Based on the United Nations High Commissioner for Refugees (UNHCR) estimates, more than 7 million people fled the country as of 5 September 2022. In this context, it is extremely important to anticipate where these people are moving so that national to local authorities can better manage challenges related to their reception and integration.

This work shows how innovative data from social media can provide useful insights on conflict-induced migration flows. In particular, we explore the potential of Facebook Social Connectedness Index (SCI) for predicting migration flows in the context of the war in Ukraine, building on previous research findings that the presence of a diaspora network is one of the major migration drivers. To do so, we first evaluate the relationship between the Ukrainian diaspora and the number of refugees from Ukraine registered for Temporary Protection or similar national schemes as a proxy of migratory flows into the EU. We find a very strong correlation between the two (Pearson’s $r = 0.94$, $p < 0.0001$), which indicates that the diaspora is attracting the people fleeing the war, who tend to reach their compatriots, in particular in the countries where the Ukrainian immigration was more a recent phenomenon.

Second, we compare Facebook’s Social Connectedness Index with available official data on diaspora at regional level in Europe. Our results suggest that the index, along with other readily available covariates, is a strong predictor of the Ukrainian diaspora at regional scale.

Finally, we discuss the potential of Facebook’s Social Connectedness Index to provide timely and spatially detailed information on human diaspora for those countries where this information might be missing or outdated, and to complement official statistics for fast policy response during conflicts.

1 Introduction

The war in Ukraine has caused the largest displacement of people in Europe since the World War II.¹ Ukraine borders four EU Member States and its citizens do not need a visa to enter the EU, being able to freely move in the Schengen area for up to 90 days (Council of European Union, 2001).² The response of the EU to those fleeing the war has been unprecedented with the activation of the Temporary Protection Directive.³ This provides a stable and relatively durable legal status which

¹Statement by Osnat Lubrani, UN Resident & Humanitarian Coordinator in Ukraine. <https://ukraine.un.org/en/175836-war-has-caused-fastest-and-largest-displacement-people-europe-world-war-ii>.

²https://eu-solidarity-ukraine.ec.europa.eu/information-people-fleeing-war-ukraine_en. Last accessed the 23 November 2022.

³https://home-affairs.ec.europa.eu/policies/migration-and-asylum/common-european-asylum-system/temporary-protection_en. Last accessed the 23 November 2022.

includes an immediate access to a set of rights, including housing, education, labour. In such a context, where large movements in a free mobility area are coupled with the immediate access to services, it is important for the authorities to understand what is driving the flow from Ukraine so to anticipate how many people are arriving in a particular country. This is true for national authorities as well as for regional authorities, which have the responsibility of guaranteeing the access to the services under their competence.

The presence of a diaspora community is one of the strongest drivers of migration. This is called “network effect” and is valid for labor migration as well as for movements of people seeking international protection (Migali et al., 2018). In addition, other important factors can push or pull the migratory flows. The income per capita gap between origin and destination (Ortega and Peri, 2009), the wage differentials between sending country and receiving country (Sjaastad, 1962; Harris and Todaro, 1970), and the share of a common language (Lanati and Venturini, 2021) are only some examples of such drivers. These are not included in this work since the aim is to focus solely on understanding the power of the diaspora on predicting the trajectories of the people displaced in the European Union (EU) due to the war in Ukraine.

The Ukrainian diaspora is particularly large in the European Union and even before the war Ukrainians were already among the top ten nationalities of non-EU born residents.⁴ This paper aims to understand how the Ukrainian diaspora might be driving the movement of people within the EU and what added value can innovative data have.

The International Organization for Migration (IOM) defines the diaspora as the group of

*migrants or descendants of migrants whose identity and sense of belonging, either real or symbolic, have been shaped by their migration experience and background. They maintain links with their homelands, and to each other, based on a shared sense of history, identity, or mutual experiences in the destination country.*⁵

The diaspora of a country X in a country Y is often measured by the presence of people born in the country X and residing in the country Y . However, in the case of the Ukrainian diaspora, the citizenship criterion is preferred to the criterion of the country of birth to account for the internal mobility occurred within the Soviet Union (Vakhitova and Fihel, 2020). In line with this, we use the term “Ukrainian diaspora” and “Ukrainian stocks” to refer to Ukrainian citizens living outside Ukraine throughout the paper. This operationalisation is used in the literature (Beine et al., 2011), however it presents some challenges. First, it does not capture short term movements of people who may not become settled migrants in the destination country but move across the origin and destination countries and have built a transitional network. This is for instance particularly relevant for Ukrainians who, thanks to their visa-free status and liberal labour migration schemes, are among the top nationalities for short-term (*i.e.* less than 12 months) residence permits in the EU.⁶ Secondly, it does not include second-generation immigrants, *i.e.* people born in the country or with the citizenship of the country of residence with at least one foreign-born parent. Thirdly, it does not capture whether the persons maintain a link with the homeland and with each other. Despite these limitations, data on the stock of foreign born or foreign citizens are to date the best available large scale comparable dataset to measure the diaspora and will be also used in this paper.

Data on the stock of migrants, however, do not always have the required spatial granularity, nor are produced with high frequency. The review of Bosco et al. (2022) has shown that innovative data can offer a great geographic and temporal granularity, a (near-) real time availability, and an extensive coverage suitable for more immediate international comparisons. An innovative way to measure a country’s diaspora is using social media data, for instance the Facebook’s ties. In particular, Facebook’s SCI measures friendship ties of active users to assess the strength of a connection between two areas.

In this paper, we analyse Facebook’s Social Connectedness Index and traditional measures of the diaspora side-by-side. Concretely, we aim to (i) understand the role of the diaspora in attracting the migratory flows during the war in Ukraine, and (ii) test the potential of Facebook’s SCI to predict the stock of Ukrainians in the EU at detailed spatial resolution.

⁴According to data from “Immigration by age group, sex and citizenship”, accessible at https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=migr_immictz.

⁵IOM Glossary 2019. https://publications.iom.int/system/files/pdf/iml_34_glossary.pdf.

⁶For instance, according to the migr_resfirst series available at <https://ec.europa.eu/eurostat/web/migration-asylum/managed-migration/database>, they were the largest recipient of such permits in 2021.

The rest of this paper is structured as follows: Section 2 examines the link between the actual flows of people fleeing Ukraine towards the EU and the Ukrainian diaspora in the 27 Member States; Section 3 analyses the relationship between Facebook’s SCI, and the Ukrainian stocks; finally, Section 4 discusses the results and concludes.

2 Analysis of the driving power of the diaspora for the refugee flows from Ukraine

The diaspora is one of the major drivers of migration (Migali et al., 2018). In this section, we assess the power of the diaspora in driving the movements from Ukraine to the EU. The diaspora is here measured as the stock of residents with Ukrainian citizenship.

2.1 Description of the data

To quantify the movements from Ukraine we use data on the refugees from Ukraine registered for Temporary Protection or similar national protection scheme.⁷ This indicator has some limitations as it does not perfectly reflect the number of Ukrainians fleeing the war who are present in the country: some people may have not registered for temporary protection (yet), hoping to return to Ukraine soon or to apply for another permit; or may have registered but moved to another country or come back to Ukraine. Yet, registrations for protection are the best proxy currently available. Data on refugees from Ukraine recorded across Europe account for people who have not registered yet; however, they present even more important limitations than the protection registrations: these data do not account for intermediate destinations, thus showing inflated counts (also known as double counts) in transit countries, especially those neighbouring Ukraine. Temporary protection data should therefore give a more likely overview of the actual net flows, since a person who registers for temporary protection in one country shows an intention to settle there for some time, especially considering that a person cannot benefit from temporary protection in more than one country at the same time.⁸ Table 1 shows the difference between the number of temporary protections and recorded refugees.

⁷<https://data.unhcr.org/en/situations/ukraine>.

⁸https://eu-solidarity-ukraine.ec.europa.eu/information-people-fleeing-war-ukraine_en#paragraph_314.

Country	Recorded refugees	Temporary Protection
Austria	78158	78158
Belgium	52245	52870
Bulgaria	129437	86722
Croatia	16828	16829
Cyprus	14989	13113
Czech Republic	412959	413121
Denmark	31000	33000
Estonia	32077	50491
Finland	35240	35240
France	96520	96520
Germany	670000	971000
Greece	18363	18363
Hungary	28289	28289
Ireland	48672	45074
Italy	150261	159968
Latvia	37496	36449
Lithuania	62444	62444
Luxembourg	6263	6263
Malta	1284	1373
Netherlands	68050	68050
Poland	1274130	1274130
Portugal	49623	49718
Romania	52952	84662
Slovakia	86834	87030
Slovenia	7200	7200
Spain	133820	133913
Sweden	42250	44107

Table 1: Number of recorded refugees from Ukraine across the European Union and refugees from Ukraine registered for Temporary Protection or similar national protection schemes.

Source: [UNHCR](#)

To quantify the diaspora we use the number of Ukrainian citizens residing in a specific country from Eurostat⁹, and, where not available, from UNDESA.¹⁰ These data sources provide diaspora measurements according to the country of citizenship, which is important since many Ukrainian citizens migrated to the EU from other countries that were previously part of the Soviet Union ([Vakhitova and Fihel, 2020](#)).

Each EU country was distinguished between two classes: “old” diaspora countries, and more recent (“new”) diaspora countries. Old diaspora countries are defined as those where the immigration of Ukrainians used to be higher in the past, and decreased recently; new diaspora countries are those which have received a stable or increasing number of migrants from Ukraine. The reason for distinguishing between a new and an old diaspora is that connections with the country of origin may deteriorate over time, so that people who migrated long time ago may have a smaller network in the country of origin than people who migrated only recently, and may thus be less of a driver for current movements. To distinguish between the new and the old diaspora we looked at the historical annual migration flows. We used data on long-term Ukrainian immigrants by age group, sex, and citizenship arriving into the reporting country data from Eurostat.¹¹ Where these data were not available, we used the International Migration Database of the Organisation for Economic Co-operation and Development (OECD)¹² or

⁹Population on 1 January by age group, sex and citizenship, accessible at https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=migr_pop1ctz&lang=en.

¹⁰International Migrant Stock. <https://www.un.org/development/desa/pd/content/international-migrant-stock>.

¹¹Immigration by age group, sex and citizenship, accessible at https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=migr_imm1ctz.

¹²International Migration Database, accessible at <https://stats.oecd.org/>.

the residence permits issued for at least 12 months from Eurostat.¹³ For this analysis, we selected the period 2008-2019, since 2008 is the first year for which data are complete. We excluded the year 2020 from the analysis since the COVID-19 pandemic would likely bias the final results.

2.2 Results

We split the historical Ukrainian immigration time series on year 2013 to obtain two main periods, each six years long: 2008-2013, and 2014-2019. We then compared the number of Ukrainian immigrants in each country before and after 2013, and classified each country as having a “new diaspora” if (i) the number of immigrants was greater in the 2014-2019 compared to 2008-2013, and (ii) if the general trend of the time series was found to increase in the last period, otherwise they were classified as having an “old diaspora”. Figure 1 shows the data used to classify each country.

In Figure 1 we have highlighted the countries classified as having an “old diaspora” with a blue box. These are Cyprus, Greece, and Italy.

To assess whether the diaspora drives the current movements from Ukraine we correlated the stock of Ukrainian citizens residing in the Member States before the war with the protection registrations from people fleeing Ukraine. The vast majority of these are Ukrainian citizens.¹⁴ The correlation coefficient is already high considering all the countries, regardless of whether the Member State is an old or a recent destination country (Pearson’s $r = 0.80$, $p < 0.0001$). When we focus on the countries with a new diaspora alone, the correlation coefficient increases to 0.94.

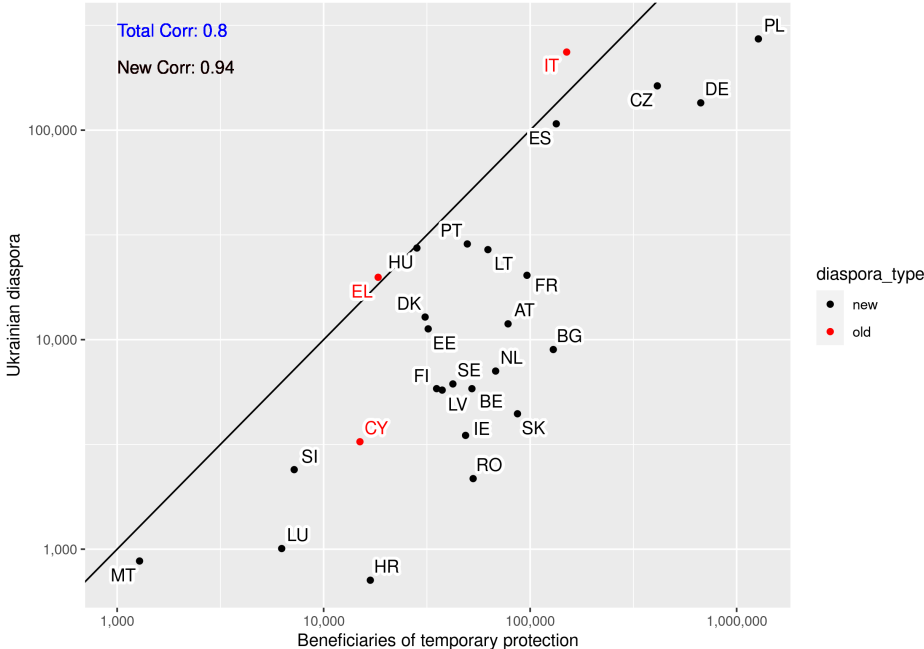


Figure 2: Number of temporary protection versus stock of Ukrainian citizens residing in the Member States before the war. The solid lines represent the bisector. Axes in logarithmic scale.

In Figure 2, we can see that amongst the countries defined as having an “old” diaspora (shown in red), Italy plays an important role, since it is the country with the largest stock of Ukrainian citizens after Poland¹⁵, but the ratio between its temporary protection registrations and the stock of Ukrainian citizens living there is relatively small compared to the other top five diaspora countries. This can be clearly seen in the scatter plot of Figure 2, where Italy lies above the bisector line whereas the rest of the countries is below. The only exception is Greece, another country classified as having an “old diaspora” type. This is an indication that despite having a strong historical Ukrainian diaspora, Italy

¹³First permits by reason, length of validity and citizenship, accessible at https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=migr_resfirst&lang=en.

¹⁴https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=migr_asytspm.

¹⁵According to data from EUROSTAT, and UNDESA.

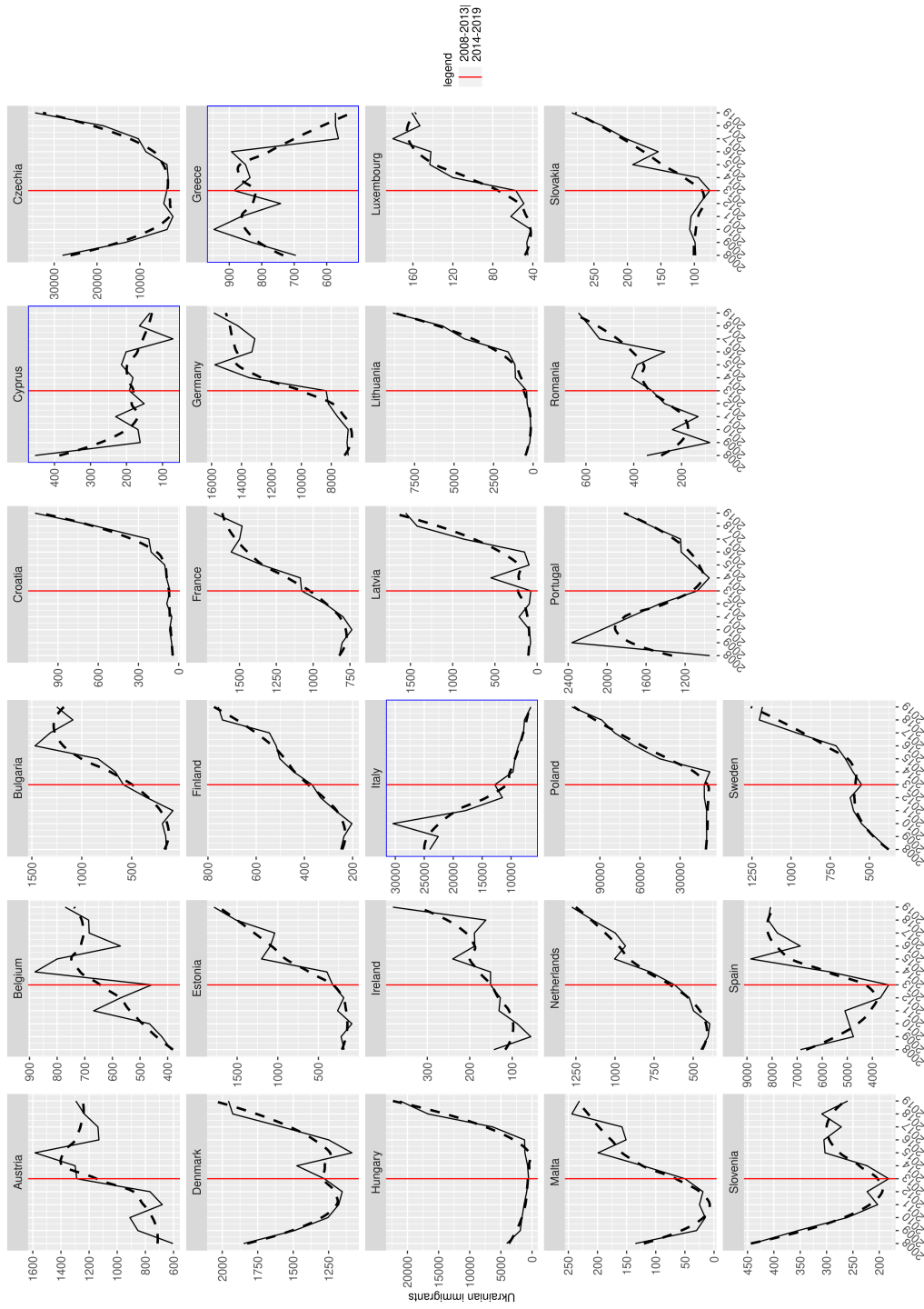


Figure 1: Number of Ukrainian immigrants by Member State during 2008-2019 (yearly data). The dashed line represents a smoothed version of the time series to better show the patterns. The vertical red line divides the period into two sub-periods (2008-2013, 2014-2019). Blue boxes highlight the countries classified as “old” diaspora.

and Greece seem less attractive for refugees than other countries with a more recent diaspora. This is likely related to other factors as well, such as economic, social, and political ones.

3 Analysis of the relationship between Facebook’s Social Connectedness Index and the Ukrainian diaspora

The findings of Section 2.1 show that the diaspora can be used to anticipate the flows. This is true at national level and reasonably also at a more granular level. While data on residents by citizenship at regional level are provided for some Member States by the national statistical offices, we do not have a harmonised and complete dataset at this spatial granularity. In this section we therefore assess the potential of Facebook’s SCI to fill this gap.

3.1 Description of the data

We collected diaspora data at regional level (*i.e.* NUTS-3) from various national statistical offices, focusing on those EU countries where the Ukrainian diaspora is the largest: Italy, Spain, Germany, Czechia, and Portugal. Unfortunately, we were not able to include data for Poland, the country with the largest stock of Ukrainian citizens in the European Union, since these were not available at the same spatial resolution as the other countries and Facebook’s SCI.

We used Facebook’s SCI¹⁶ (Bailey et al., 2020) before the start of the conflict in Ukraine and the total population of the hosting countries¹⁷ to evaluate their potential to predict human diaspora - the response variable of our analysis. Facebook’s SCI uses a snapshot of Facebook users and their friendship networks to measure the intensity of connectedness between locations in a specific time. Locations are assigned to Facebook users based on their information and activity on Facebook, including the stated city on their Facebook profile, and device and connection information. These locations do not imply a user is a citizen of that country. Facebook’s SCI between two locations i and j is calculated as:

$$Social\ Connectedness_{ij} = \frac{FB\ Connections_{ij}}{FB\ users_i * FB\ users_j} \quad (1)$$

$FB\ users_i$ and $FB\ users_j$ are the number of Facebook users in locations i and j , and the quantity $FB\ Connections_{ij}$ is the number of Facebook friendship connections between the two. The indicator $Social\ Connectedness_{ij}$, therefore, is proportional to the likelihood that a given Facebook user in location i is friend on Facebook with a given user in location j .

Figure 3 shows the connections between Ukraine (considering the whole country), and each NUTS-3 area within the European Union.

¹⁶<https://dataforgood.facebook.com/dfg/tools/social-connectedness-index>.

¹⁷Population on 1 January by broad age group, sex and NUTS-3 region. https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo_r_pjanaggr3.

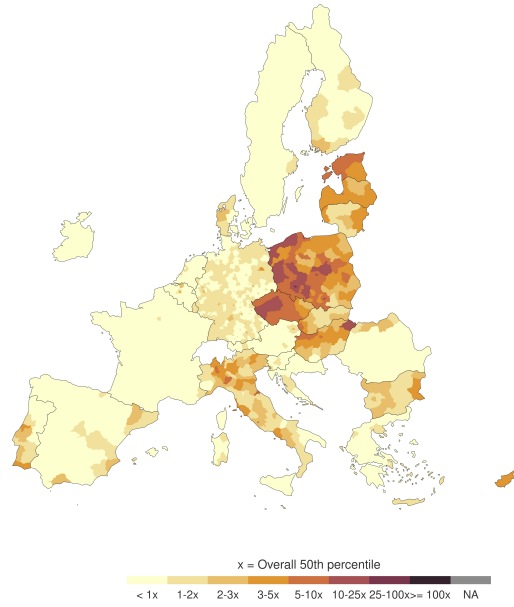


Figure 3: Facebook Social Connectedness Index (Ukraine to EU27 NUTS-3). This map is for illustration purposes only. The boundaries and names shown and the designations used on this map do not imply official endorsement or acceptance by the International Organization for Migration.

Source: Eurostat, ©EuroGeographics for the administrative boundaries

3.2 Results

To test the hypothesis that the virtual connections (*i.e.* Facebook friendship ties) between Ukraine and EU regions resemble the Ukrainian diaspora within the Union, we examined the correlation between Facebook SCI and Ukrainian diaspora at NUTS-3 level, however the strength of this correlation varies significantly among countries (see Table 2). Therefore, we introduced the total population and the geographic position of each region as additional covariates in the analysis, as intuitively these are potentially linked to the diaspora. Indeed, the selected variables proved to be very good predictors for the diaspora.

Table 2: Correlation coefficients of Facebook SCI and Ukrainian diaspora by country.

country	corr_coeff ($p < 0.0001$)
CZE	0.87
DEU	0.47
ESP	0.65
ITA	0.43
PRT	0.54

The Ukrainian diaspora, Facebook’s SCI, and the total population datasets all present a right-skewed distribution, meaning that each of these dataset has very few observations with very high values (compared to the mean), and these represent the capitals or other important areas (Figure 4).

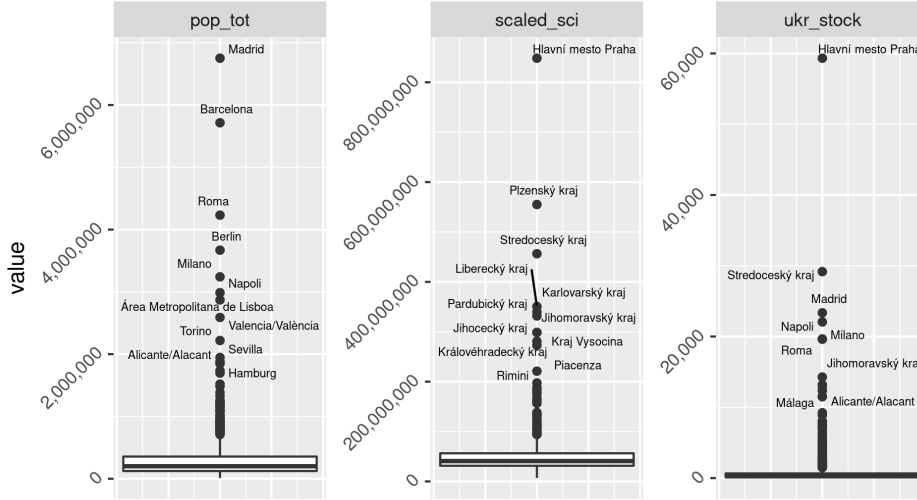


Figure 4: Distribution of total population (left), Facebook SCI (middle), and Ukrainian diaspora (right).

In other words, usually the most populated NUTS-3 are also those with a larger diaspora and/or a higher level of social connectedness. Even though these regions are very few, they have a major role in driving the migratory flows. This is why we introduced a weight variable in the analysis based on the interaction between Facebook’s SCI and the total population. We derived our weights as the product between the two variables normalised between 0 and 1 through a min-max approach.

$$W = \tilde{P} \cdot \tilde{S} \quad (2)$$

In Equation 2, \tilde{P} and \tilde{S} represent the normalized vectors of total population and Facebook SCI respectively, and W is the vector of the resulting weights.

Finally, we used the following model to evaluate the predictive power of the chosen variables:

$$\log(ukr_stock) = \alpha + \beta_1 \log(scaled_sci) + \beta_2 \log(pop_tot) + \beta_3 latitude + \epsilon \quad (3)$$

We randomly selected 70% of the total 594 observations to build a sample to train the model, and we used the remaining set to evaluate its performance. The explained variance (R^2 , calculated on the validation set using the formula found in [Bosco et al. \(2017\)](#)) is 0.92, and all predictors have a highly significant p-value (< 0.01). Figure 5 shows the scatter plot of the observed diaspora values versus the ones predicted by the model, where each color represents one of the countries available in the dataset.

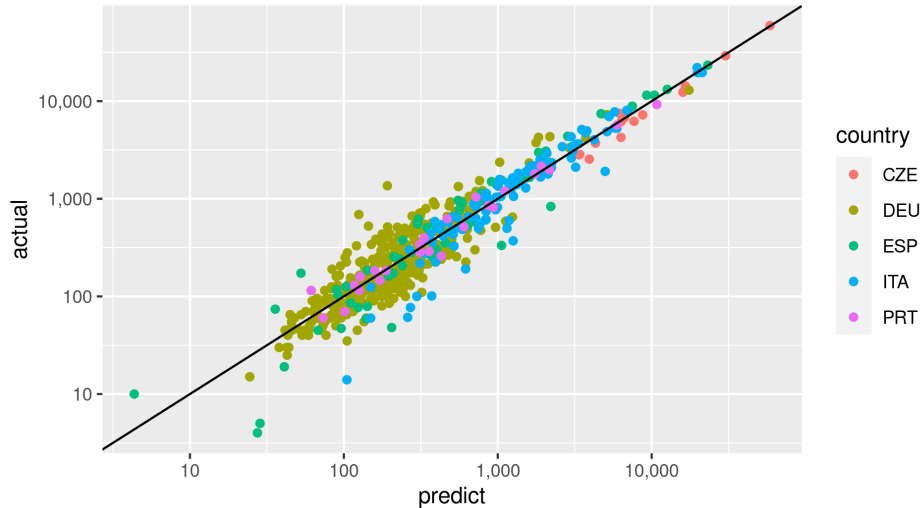


Figure 5: Scatter plot of the observed diaspora values (*actual* on the *Y* axis) versus the ones predicted by our model (*pred* on the *X* axis) coloured by country. Axes in logarithmic scale.

4 Discussion and conclusions

The war in Ukraine has provoked the largest human displacement in Europe in recent years. As a response, the EU has activated the Temporary Protection Directive, which provides for immediate access to a variety of services to its beneficiaries. In this context, it becomes paramount for the authorities to be prepared and to anticipate how many people will be arriving. This is true for national as well as for local authorities, who are responsible for several of the services covered by the Directive.

The literature shows that the diaspora is a key driver of migration via a “network effect” (Migali et al., 2018). In this paper, we first verified the power of the diaspora to anticipate the population movements of Ukrainians following the Russian invasion. Secondly we explored the potential of innovative data, more specifically Facebook’s SCI, to measure the diaspora.

The main results of this analysis are the following.

First, the diaspora is confirmed to be an important driver of migration, especially when it is the results of migration movements that increased more recently or did not decrease over time (Section 2.1). The correlation between the number of Ukrainian citizens residing in EU countries (as a proxy for the diaspora) and the number of registrations for temporary protection or similar national schemes (as a proxy for migration movements) is high and significant, especially when ‘old’ migration countries are excluded. This is consistent with the hypothesis that people who migrated long time ago may have a smaller or weaker network than those who migrated more recently.

Secondly, Facebook’s SCI, along with data on total population and the geographical position of the regions, can be used to measure the diaspora at a granular spatial resolution (NUTS-3). This is useful as, while the need to anticipate the arrivals is important also at regional level, spatially detailed data on the diaspora are often not available in a systematic manner. Innovative data can be used to fill this gap.

This paper paves the way for potential future research. In particular, the use of Facebook’s SCI to measure the diaspora can be further explored. Besides more granular information at spatial level, Facebook’s SCI can provide information also at high frequency in time. This can be used to more timely monitor and anticipate changes in the migratory flows occurring before the publication of official statistics.

Moreover, Facebook’s SCI can be used in complementarity with the data on migration stock to provide for a more comprehensive operationalisation of the diaspora. While widely used in the literature, the migrants stock cannot capture all the defining elements of the diaspora, in particular whether a meaningful link with the country of origin exists. Facebook’s SCI can help attest the presence of such link, although it does not tell us whether it indicates a “shared sense of history, identity” and

whether it has been forged by a migration experience occurred in the current or previous generations. In addition, other important factors can push or pull the migratory flows. The income per capita gap between origin and destination, wage differentials between sending country and receiving country, and the share of a common language are only some examples of such drivers. Therefore, our future research will focus on the inclusion of the Facebook’s innovative indicator in a more complex model to predict migration movements, in addition to data on the diaspora and other traditional indicators, also going beyond the Ukrainian case.

References

- Bailey, M., D. Johnston, T. Kuchler, D. Russel, B. State, and J. Stroebel (2020). The determinants of social connectedness in europe. In *Social Informatics*. Springer International Publishing.
- Beine, M., F. Docquier, and Çağlar Özden (2011). Diasporas. *Journal of Development Economics* 95(1), 30–41. Symposium on Globalization and Brain Drain.
- Bosco, C., V. Alegana, T. Bird, C. Pezzulo, L. Bengtsson, A. Sorichetta, J. Steele, G. Hornby, C. Ruktanonchai, N. Ruktanonchai, et al. (2017). Exploring the high-resolution mapping of gender-disaggregated development indicators. *Journal of The Royal Society Interface* 14(129), 20160825.
- Bosco, C., S. Grubanov-Boskovic, S. M. Iacus, U. Minora, F. Sermi, and S. Spyrtatos (2022). Data innovation in demography, migration and human mobility. Technical Report EUR 29333 EN, European Commission - Joint Research Centre, Luxembourg (Luxembourg).
- Council of European Union (2001). Council directive 2001/55/ec of 20 july 2001 on minimum standards for giving temporary protection in the event of a mass influx of displaced persons and on measures promoting a balance of efforts between member states in receiving such persons and bearing the consequences thereof.
<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32001L0055>.
- Harris, J. R. and M. P. Todaro (1970). Migration, unemployment and development: A two-sector analysis. *The American Economic Review* 60(1), 126–142.
- Lanati, M. and A. Venturini (2021). Cultural change and the migration choice. *Review of World Economics* 157(4), 799–852.
- Migali, S., F. Natale, G. Tintori, S. Kalantaryan, S. Grubanov-Boskovic, M. Scipioni, F. Farinosi, C. Cattaneo, B. Benandi, M. Follador, G. Bidoglio, T. Barbas, and S. McMahon (2018). International migration drivers. Technical Report EUR 29333 EN, European Commission - Joint Research Centre, Luxembourg (Luxembourg).
- Ortega, F. and G. Peri (2009, April). The causes and effects of international migrations: Evidence from oecd countries 1980-2005. Working Paper 14833, National Bureau of Economic Research.
- Sjaastad, L. A. (1962). The costs and returns of human migration. *Journal of political Economy* 70(5, Part 2), 80–93.
- Vakhitova, H. and A. Fihel (2020). International migration from ukraine: Will trends increase or go into reverse? *Central and Eastern European Migration Review* 9, 125–141.