European Commission

JRC TECHNICAL REPORT

Social connections and remittance flows: an exploratory analysis of Facebook data

Kalantaryan, S
McMahon, S
Spyratos, S
Iacus, S

2022

Joint Research Centre

EUR 31118 EN

Contact information
Name: Sona Kalantaryan
Email: Sona.KALANTARYAN@ec.europa.eu

EU Science Hub
https://ec.europa.eu/jrc

JRC129744

EUR 31118 EN

PDF ISBN 978-92-76-53484-6 ISSN 1831-9424 doi:10.2760/318892

Luxembourg: Publications Office of the European Union, 2022

How to cite this report: Kalantaryan, S; McMahon, S; Spyratos, S and Iacus, S *Social connections and remittance flows: an exploratory analysis of Facebook data*, EUR 31118 EN, Publications Office of the European Union, Luxembourg, 2022, ISBN 978-92-76-53484-6, doi:10.2760/318892, JRC129744.

Contents

Acknowledgements 1
Abstract 2
Summary 3
1 Introduction 4
2 The data context 6
2.1 Remittance statistics 6
2.2 Exploring migrant connections 7
2.2.1 Facebook connections 8
3 Method and results 10
3.1 Summary statistics 10
3.2 Empirical strategy 11
3.3 Regression results 12
3.4 Robustness checks 15
4 Conclusions and recommendations 16
References 17
List of tables 20
Annexes 21
Annex 1. Balance of payment items reported by Eurostat 21
Annex 2. Data sources used in the study 22
Annex 3. Robustness checks 23

Acknowledgements

The authors thank colleagues from the European Commission for their support and feedback throughout the research process. Particular appreciation goes to Iacopo Viciani (INTPA), Heliodoro Temprano Arroyo (ECFIN) and Robert Leisch (ESTAT) for their comments on a previous draft. All content and potential errors are the responsibility of the authors.

Authors

Sona Kalantaryan (Joint Research Centre)

Simon McMahon (Joint Research Centre)

Spyridon Spyratos (Joint Research Centre)

Stefano Iacus (Harvard University)

Abstract

Remittances are a major source of finance for countries around the world and are increasingly important for sustainable development. However, the available data estimating international remittance flows is often criticised. This report makes a novel contribution to understanding the factors which influence the volume of bilateral remittance flows, and in doing so highlights a potential opportunity for the refinement of international remittance flow estimates. Specifically, the study uses innovative data from Facebook's Social Connectedness Index to explain bilateral remittance flows. This data source is positively associated with remittance flows and the findings are robust and statistically significant. These findings have two key implications. Firstly, they show that, alongside data on migrant stocks, economic and geographic variables the connections between migrants and their countries and communities of origin make an important contribution to explaining bilateral remittance flows. Secondly, they illustrate how innovative data from non-traditional sources can provide new opportunities for analysing and estimating bilateral remittances.

Summary

Migrant remittances are a major source of finance for countries and households around the world. The European Union supports several initiatives which aim to lower the cost of remitting, enable people to send remittances and support their development impact in receiving countries. Data on remittance flows is vital for these and other policies and programmes, in particular for Sustainable Development Goal target 17.3.2, which aims to mobilise additional financial resources for developing countries.

However, the currently available international data on remittance flows has been criticised. Refining the way that remittance flow data is calculated would be beneficial for planning, implementing and monitoring international and national policies and programmes supporting sustainable development around the world.

This Technical Report explores the explanatory potential of data on social connections between migrants and their communities and countries of origin. In doing so, it makes a novel contribution to understanding the factors which influence the volume of bilateral remittance flows and highlights an opportunity for the refinement of international statistics. Specifically, the report has two key findings:

First, it shows that the incorporation of data on the social connections between migrants and their countries and communities of origin into current calculations could potentially refine the accuracy of bilateral remittance flow estimates. The study finds that data on international social connections between migrants and their countries and communities of origin can strengthen statistical models explaining bilateral remittance flows. These findings are robust and statistically significant, although it should be noted that they demonstrate an association, not causality.

Second, the report highlights how innovative data can provide an opportunity for analysing and estimating remittance flows on a global level. This is confirmed by an analysis of data from Facebook's Social Connectedness Index, which provides information on connections between people in different countries (via Facebook) on a scale which has not been previously possible with other research approaches. When Facebook data on connections between people in different countries is examined alongside data on the migrant population in a given country, economic and geographic variables, it leads to stronger explanations of the scale of bilateral remittance flows.

In light of these findings, the report proposes further reflection and exploration of the ways in which innovative data on connections between migrants and their countries and communities of origin can explain the scale of remittance flows and feed into future estimate calculations. There are numerous potential avenues to explore, including:

- Global air traffic: Air passenger data records individual journeys between two airports by airplane. The number of journeys made between countries can be an indicator of the connections between people living and working in each, including (but not only) migrants and diasporas. Air traffic data can be accessed directly from private companies such as IATA and Sabre. Proxy data can also be accessed through Eurostat, within annual International Trade in Services Statistics and the JRC has produced an open-source dataset on mobility data including tourism and air traffic passenger data.

- Mobile money-transfer services: mobile money transfer services are expanding rapidly, and could potentially provide a new data source on the volume and direction of flows between countries. There are already examples of this data being analysed in studies in France and Rwanda.

- Google trends: geo-referenced online search data from Google Trends can show the prevalence of certain keywords in different places. This could potentially be used to explore the extent to which people who reside in one country search for information about the situation in another. These searches may indicate an ongoing connection between a migrant and their country of origin, as they seek to keep informed about life there.

The recently formed International Working Group to Improve Data on Remittance Flows, led by the World Bank under the umbrella of the KNOMAD network, would be an ideal venue for these explorations to take place.

1 Introduction

Migrant remittances are a major source of finance for countries and households around the world. In 2020, remittance inflows to low and middle income countries were worth an estimated US$540 billion, surpassing aggregate Foreign Direct Investment (FDI) or Overseas Development Assistance (ODA) flows (McMahon and Kalantaryan 2021).[1] Data on these remittance flows is vital for international and national policies and programmes. The World Bank-led KNOMAD network has stated that 'accurate and timely data on remittance flows are necessary to enable decision-makers to enhance evidence-based policies and support the achievement of (the) sustainable development goals' (KNOMAD 2022). Data on remittance flows is associated with Sustainable Development Goal (SDG) target 17.3.2 which aims to mobilise additional financial resources for developing countries, and can be significant for a country's access to international capital markets by impacting on macro-economic calculations of creditworthiness (Avendaño et al 2009).

However, international statistics on aggregate remittance flows have been criticised (Alvarez et al 2015; KNOMAD 2017, 2022; Lubambu 2014; OECD 2006). There are gaps in the available data, such as when countries do not record remittance-related financial flows or when countries do not record the information in the same way as each other (KNOMAD 2017; Lubambu 2014). National statistics tend to show the total scale of flows in and out of a country, but not bilateral flows between individual sending and receiving countries (Leuth and Ruiz-Arranz 2008). Informal remittance flows, such as those sent through *hawala* or by cash-in-hand are not registered in banking systems and remain invisible for national statistics.[2] These are just a few of the issues to be addressed by a newly formed International Working Group to Improve Data on Remittance Flows, led by the World Bank under the umbrella of the KNOMAD network (KNOMAD 2022).

In this report we make a novel contribution to understanding the factors which influence the volume of bilateral remittance flows, and in doing so highlight a potential direction for the refinement of international statistics. Current estimates of bilateral remittance flows draw primarily on a combination of migrant stock data, national current account and capital transfer data and socio-economic indicators. Our analysis shows that, when examined alongside economic and geographic variables, international social connections between migrants and their countries and communities of origin are positively associated with bilateral remittance flows. This is confirmed by an analysis of innovative data from Facebook's Social Connectedness Index.

Our findings have two main implications for the production of data estimating remittance flows. Firstly, they show that the incorporation of data on the social connections between migrants and their countries and communities of origin into current calculations could potentially refine the accuracy of bilateral remittance flow estimates. Secondly, they highlight an opportunity provided by innovative data for analysing and estimating remittance flows on a global level. We recommend further reflection and exploration of the ways in which innovative data on connections between migrants and their countries and communities of origin can feed into estimates of remittance flows in the context of the aforementioned International Working Group.

Better data on remittance flows will be beneficial for a range of European, international and national policies and programmes. Indeed, the SDGs call for increasing significantly the availability of high-quality, timely and reliable data (Target 17.18) and supporting statistical capacity-building in developing countries (Target 17.19). Eurostat's Director General Mariana Kotseva has also underlined how data and statistics are 'a key tool for governments and policy makers to measure progress towards development goals and provide information about the effectiveness of policies and programmes' (Eurostat 2021). Regarding remittances in particular, the European Union (EU) supports several initiatives which aim to lower the cost of remitting, enable people to send remittances and support their development impact in receiving countries.[3] Having good data on remittance flows is important to plan, implement and monitor these and other initiatives around the world.

[1] For the latest estimates of remittance flows, see World Bank Migration and Remittances Data available at https://www.worldbank.org/en/topic/migrationremittancesdiasporaissues/brief/migration-remittances-data, accessed 1st June 2022

[2] In theory these transactions should be estimated and included in national Accounts and Balance of Payments whenever there is a transfer between a resident and a non-resident household. The fact that these transactions are not directly observable is no reason for excluding them (similar issues apply to e.g. import of illegal goods).

[3] In Africa the EU has financed the 'African Postal Financial Services Initiative supporting and enabling African post offices to provide financial services in Benin, Ghana, Madagascar and Senegal, and the 'Maximizing the Impact of Global Remittances in Rural Areas' (MIGRRA) programme also focusing on lowering remittance costs on the Kenya-Uganda corridor (both led by IFAD). The 'PRIME Africa' initiative brings together public and private stakeholders in 9 African

The report is structured around the three chapters which follow. The first examines the current data context related to remittances, summarising existing approaches to explaining and calculating remittance flows and introducing innovative data from Facebook. The next reviews the statistical models and analyses employed in this particular study and presents the obtained results. The final chapter summarises the main findings.

countries (Cabo Verde, Ethiopia, Ghana, The Gambia, Kenya, Morocco, Senegal, South Africa and Uganda) to make remittances cheaper and faster with the support of National Remittance Task Forces. Last but not least, the InclusiFI project aims to enable migrants and diasporas to contribute to development by investing in micro-, small and medium-sized enterprises

2 The data context

Existing research on remittance flows is diverse and heterogeneous, with remittances being conceived and measured in a range of different ways (IMF 2009a, 2009b). The definition of remittances which underpins most international statistics is that set out in the International Monetary Fund (IMF) Balance of Payments and International Investment Position Manual Sixth Edition (BPM6). This describes remittances as 'household income from foreign economies arising mainly from the temporary or permanent movement of people to those economies' (IMF 2009b: 272). It is the definition that we use throughout this study. The sections below outline the composition of international statistics which are based on this definition and the potential contribution of data on social connections between migrants and their country of origin.

2.1 Remittance statistics

The BPM6 delineates the standard components of remittances within the balance of payments framework. These are defined as (see also Table 1 below); (a) personal transfers, which is understood as 'all current transfers in cash or in kind made or received by resident households to or from nonresident households … (including) all current transfers between resident and nonresident individuals' (b) compensation of employees, which is understood as 'the income of border, seasonal, and other short-term workers who are employed in an economy where they are not resident and of residents employed by non-resident entities', and (c-f) supplementary items, which include capital transfers, capital transfers between households, social benefits and others (although the Manual states that their 'compilation and dissemination is encouraged but voluntary' and others note that they are rarely reported, see IMF 2009b; KNOMAD 2017). Aggregating these components gives three configurations of remittances:

> *Personal remittances (R1)* - personal transfers plus compensation of employees (less taxes and social contributions paid by non-resident workers in the economy of employment, less transport and travel expenditures related to working abroad). In other words, it includes all household-to-household transfers and the net earnings of non-resident workers[4]. (BPM6, A5.10)
>
> *Total remittances (R2)* - the sum of personal remittances and social benefits. (BPM6, A5.14).
>
> *Total remittances and transfers to Non-profit institutions serving households (NPISHs) (R3)* - total remittances and both current and capital transfers to NPISHs from any sector of the sending economy. It, therefore, includes donations, in cash or kind, from government and enterprise sectors to charitable organizations in another economy[5]. (BPM6, A5.15).

Table 1. Tabular presentation of the definitions of remittances

Total remittances and transfers to NPISHs (a+b+c+d+e+f)					
Total remittances (a+b+c+d)					
Personal remittances (a+b+c)					
a	b	c	d	e	f
Personal Transfers (part of current transfers)	Compensation of employees less taxes, social contributions, transport and travel	Capital transfers between households	Social benefits	Current transfers to NPISHs	Capital transfers to NPISHs

Note: Personal transfers is a standard item; other items are supplementary.
Source: BPM6, IMF (2009)

[4] It is worth mentioning that the various transactions (travel expenditures, compensation of employees, personal transfers and capital transfers) are recorded under different current account items and have an impact on various core macro-economic indicators like GDP, GNI, disposable income and net borrowing net lending.

[5] The compensation of employees (for non-resident migrant workers) and personal transfers (for resident migrant workers) between two countries have to be reported according to the recording standards of the BPM6. This distinction has an impact on the recordings in National Accounts (GDP, GNI vs. DPI) and Balance of Payments.

Data based on the balance of payments provide the foundation for the main international estimates of remittance flows. The includes the data for SDG Indicator 17.3.2 (the volume of remittances as a proportion of total GDP),[6] estimates and forecasts of current and future remittance flows[7] and the Bilateral Remittance Matrix,[8] all of which are produced by the World Bank. The World Bank's approach to calculating remittance flows complements balance of payment data and central bank reports with World Bank staff calculations. As explained in the Migration and Development Brief 28 (KNOMAD 2017), this involves comparing two different projections. One is 'based on partial quarterly or monthly year-to-year growth rates (usually based on data from the central bank or national statistical office), and applying that growth rate to the previous year for which the data are available'. The other is based on a different calculation using estimates based on two main datasets: (i) UNDESA's Bilateral Migration Matrix which provides estimates of immigrant stocks in all countries disaggregated by countries of origin and (ii) balance of payments remittance data (i.e. the sum of compensation of employees and personal transfers). The estimated bilateral flows are calculated by distributing total remittance inflows according to weights calculated using emigrant stocks from the Bilateral Migration Matrix and adjusted for per capita income in sending and receiving countries. This methodology also informs the Bilateral Remittance Matrix, which is a dataset with global coverage providing estimates of remittance flows between specific countries for the period of 2010 – 2018. The dataset is widely used by researchers and policymakers.

Eurostat also provides data on remittances sent and received by EU Member States and selected non EU countries, based on the new standard BPM6. However, this was introduced in June 2014[9] and for the moment only the standard components of remittances are recorded. According to a methodological note, the broader measures (R2 and R3) cannot yet be compiled by Eurostat. However, also the compilation of personal remittances (R1) is incomplete due to the lack of supplementary components, resulting from the compensation of employees being available predominantly on a gross basis, and missing cross-border capital transfers between households (Obrzut 2020)[10].

2.2 Exploring migrant connections

As described above, estimates of remittance flows are calculated by drawing mainly on data which refers to (a) reporting in balance of payments statistics ; (b) migrant stock data, and (c) other socio-economic variables such as GDP per capita (KNOMAD 2017; Lueth and Ruiz-Arranz 2008). The presence of migrants from a given country of origin residing in another country is one of the main predictors of remittance flows between them (Schiopu and Siegfried 2006; Hagen-Zanker and Siegel 2007; Lueth and Ruiz-Arranz 2008). But this does not fully capture which migrants remit and why.

In this report we underline the economic importance of interpersonal international connections between migrants and their countries and communities of origin. Research has repeatedly shown the significance of migrant networks within and between countries for international migration. This includes studies showing how diaspora networks lower the costs of migrating (e.g. Beine and Parsons 2015; Beine et al 2016) and of the importance of network effects in explaining aggregate international migration flows (Migali et al 2018). Networks have also been associated with remittances. For example, Roberts and Morris (2003) have found that stronger and broader networks of migrants and communities of origin lead to higher remittances. This is because by remitting migrants maintain connections which can enable them to participate in local employment networks. Similarly, Rosales (2020) highlights how migrant networks lead people to donate to projects in their home-town of origin, partly as a result of peer-group pressure among emigrants and partly as a way to maintain connections with the local community of origin which would be useful if they were to return in the future. Fenoll and Kuehn (2018) have also shown that migrant networks impact on the amount that people remit and how. In particular, they find that 'immigrants from above-average remitting countries remit more if they live in larger

[6] For more information, see https://unstats.un.org/sdgs/metadata/?Text=&Goal=17&Target=17.3 (accessed 1st June 2022)
[7] See World Bank Migration and Remittances Data available at https://www.worldbank.org/en/topic/migrationremittancesdiasporaissues/brief/migration-remittances-data, accessed 1st June 2022
[8] For more information, see https://www.knomad.org/data/remittances (accessed 1st June 2022)
[9] European balance of payments statistics in line with EC Regulations 184/2005 and 555/2012 on community statistics concerning balance of payments, international trade in services and foreign direct investment. Available at https://eur-lex.europa.eu/legal-content/en/TXT/?uri=celex:32005R0184, accessed 1st June 2022
[10] The authors presents four reasons why these components are missing: (i) voluntary character of supplementary components in the BPM6; (ii) relative insignificance of remittance flows; (iii) heterogeneous character of remittance transactions and (iv) bilateral asymmetries.

networks … (and) migrants in larger networks are less likely to use most expensive remittance channels.' In sum, migrants who are connected with other migrants who are originally from the same place and with people in their place of origin seem to be more likely to send remittances.

However, to date there has been a lack of data at a global scale which can measure connections between migrants and their places of origin. Existing studies have instead tended to rely on surveys or individual case studies. 'Innovative' or 'non-traditional' data sources can potentially contribute to filling this gap. Innovative data can be summarised as 'data derived from an individual's digital footprint, from sensor-enabled objects, and/or … inferred using algorithms' (Bosco et al. 2022). Their production and distribution is closely related to developments in, and the propagation of, digital technologies and their use is most common in situations of data scarcity, such as when there are gaps in the availability of other forms of statistics. Bosco et al (2022) provide an overview of opportunities and challenges related to the use of non-traditional data in demography migration and human mobility. Regarding remittances, they highlight how data from mobile phone-based money transfer services has been used to estimate the volume of remittances sent from France, and to track financial transfers in response to the 2008 earthquake in Rwanda. This study uses a specific source of innovative data on migrants' international social connections, specifically the Facebook Social Connectedness Index (SCI).

2.2.1 Facebook connections

Facebook is the world's most widespread social media platform. In April 2022, it had nearly three billion active users worldwide.[11] Data on Facebook users has already been used to estimate international migration flows and monitor displacement in the context of a national disaster (Spyratos et al. 2019; Zagheni 2020). Past migration patterns have been highly correlated with an increase in observed Facebook friendship ties (Bailey et al. 2018; Tjaden 2021).

The Social Connectedness Index (SCI) is a specific data source which measures the strength of connectedness between two geographic areas as represented by Facebook friendship ties.[12] It has been used to understand how digital social connections in Europe are related to patterns of migration, past and present political borders, geographic distance, language, and other demographic characteristics (Bailey et al. 2020). SCI is simply a measure of the density of digital social connections around the world at the national and sub-national levels. It is developed based on the total number of friendship links of Facebook users between geographic areas. Each friendship between people in different geographic areas is counted as a connection.[13] We interpret a connection as an ongoing relationship between people who have moved geographically apart from one another. As with the migrant networks described above, connections may be motivated by a range of different personal, emotional or economic reasons. They might enable family members to keep up to date with each other's lives, or include people in networks which provide employment opportunities.

In this study we use SCI to develop a country-pair specific Facebook Connection Indicator, which is calibrated to account for differences in Facebook penetration rate across countries. The departure point is the SCI which is computed as the ratio of Facebook connections between countries i and j to the product of Facebook users in the two countries as presented in Equation (1).

$$SCI_{ij} = \frac{Facebook\ Connections_{ij}}{Facebook\ Users_i\ X\ Facebook\ Users_j} \quad (1)$$

Here the two elements in the denominator are the number of Facebook users in countries i and j ($Facebook\ Users_i$ and $Facebook\ Users_j$). $Facebook\ Connections_{ij}$ is the number of Facebook friendship connections between the two locations (countries). To obtain $Facebook\ Connections_{ij}$ we follow the formula presented in Equation (2), multiplying the SCI by both the estimated number of $Facebook\ Users_i$ and $Facebook\ Users_j$.

[11] Facebook statistics and trends. Available at https://datareportal.com/essential-facebook-stats

[12] The information is provided by Meta though Data for Good initiative. For more information see https://dataforgood.facebook.com/dfg/docs/methodology-social-connectedness-index, accessed 1st June 2022

[13] The dataset is available through the Humanitarian Data Exchange repository. The reference date of the SCI data is 16 December 2020. For more information see https://dataforgood.facebook.com/dfg/tools/social-connectedness-index

$$Facebook\ Connections_{ij} = SCI_{ij}\ X\ Facebook\ Users_i\ X\ Facebook\ Users_j \quad (2)$$

However, the number of Facebook users at the country level is not publically available. Hence, we need first to estimate number of Facebook users for each country on a global scale. To do so we developed a Python script, which was used to query the Facebook Marketing Application Programming Interface (API).[14] Having estimates of Facebook users allows us to calculate bilateral Facebook connections between two countries.

At this point, we are ready to calculate our own Facebook Connections Indicator that is independent of whether Facebook is more or less popular in one country than in another, following Equation (3).

$$Facebook\ Connections\ Indicator_{ij} = \frac{Facebook\ Connections_{ij}}{Facebook\ Connections_i}\ X\ Population_i \quad (3)$$

Here $Facebook\ Connections_i$ is the sum of the total connection in the country *i* which includes both the connections to other countries and within the same country; $Population_i$ is the total population of country *i* (we used UNDESA statistics and interpolated the data for missing years). For instance, the Facebook Connections Indicator for Albanians in Italy would be calculated as the ratio between Facebook Connection between Albania - Italy to all Facebook Connections originating in Italy (within Italy plus towards the rest of the world) multiplied by the population of Italy.

[14] For more info see https://developers.facebook.com/products/marketing-api/ x\

3 Method and results

In our empirical exercise, we use data on bilateral remittances coming from the 'Personal transfers and compensation of employees' dataset provided by Eurostat. Unlike the Bilateral Remittance Matrix, the information provided by Eurostat is not calculated as a distribution of total flows through weights based on the stock of migrants and difference in income level. Rather, it comes directly from countries' Balance of Payments reporting. This allows us to use the stock of migrants and the difference in income level (GDP per capita) as explanatory variables without creating methodological concerns. On the negative side, Eurostat's data is limited in terms of geographic coverage. Actually, there are only a small number of remittance corridors for which the main variable of interest, namely ' R1- Personal Remittances' (outflow from country *i* to *j*), is effectively reported. We use debit for countries *i* and *j* to describe the outflow of remittances from a country *i* (where the migrant resides) to *j* (where he/she sends money). This is because more precise and comprehensive information coming from compensation of employees based wage tax (payslip) is available for the debit side of a country whereas for the credit side more assumptions concerning the country breakdown and the total earned gross compensation are needed. We end up having 312 observations. These observations were limited only to remittances sent and received by five EU and one non-EU (accession) country. Those countries are Chechia, Finland, Croatia, Lithuania, Romania, and Albania. We have information on the volume of remittances sent from each of these countries to 58 countries. To increase the number of observations we take the credits between countries *i* and *j* and transformed those into debit from country *j* to *i*. In this way, we reached 639 observations for R1 leaving the coverage still quite limited[15]. To attenuate this we look at three extra variables (BoP elements) reported by Eurostat and as a result, we have in total the following four variables:[16]

- R1 Personal Remittances
- R1WR Workers remittances and compensation of employees
- D752 Personal transfers (Current transfers between resident and non-resident households)[17]
- D752W Workers' remittances.

The variables originally reported in euros are converted into dollars.

3.1 Summary statistics

To explain the bilateral remittance flow between countries, we use UNDESA's Bilateral Migration Matrix, the difference in GDP per capita between the countries, the distance between countries (distance between major cities and whether the countries share a common border), and the novel variable on Facebook Connections. Table 2 presents the summary statistics of dependent and explanatory variables used in the empiric exercises presented in this study.

[15] Please, note that not all observations available in the dataset entered the sample on which the regressions were run. This is caused by missing values for other explanatory variables included in the model.

[16] Annex 1 provides an overview of all elements of Balance of Payments items available in Personal transfers and compensation of employees dataset.

[17] Obrzut (2020) states 'Personal transfers consist of all current transfers in cash or in kind between resident and non-resident households, disregarding the underlying source of income, the relationship between the households or the purpose of the transfer. Thus, the concept of personal transfers is seen to be broader than workers' remittances in the BPM5 and reflects prevailing compilation practice in some countries, where it appears difficult to differentiate between migrant transfers and donations. However, the reporting of a memo item for workers' remittances is still maintained in the BPM6. In line with its broader definition we would expect personal transfers to be higher than worker's remittances (including donations)'.

Table 2. Summary statistics

Variable	Unit	Obs.	Mean	St. Dev.	Min	Max
R1 Personal Remittances	US$ million	586	26.23	116.44	0.00	1420.55
R1WR Workers remittances and compensation of employees	US$ million	2,868	40.36	353.56	0.00	13563.93
D752 Personal transfers	US$ million	2,866	11.66	75.76	0.00	2534.25
D752W Workers' remittances	US$ million	2,621	12.63	82.94	0.00	2534.25
World Bank Bilateral Remittances (subset)	US$ million	2937	48.00	206.80	0.00	3178.56
Stock of migrants	1000	2937	15.03	79.77	0.00	2141.72
Facebook Connections Indicator	1000	2937	33.08	112.39	0.00	1916.70
Distance	km	2937	5720.36	4473.03	59.62	19335.40
Shared border	1/0	2937	0.04	0.19	0.00	1.00

Note: For all variables except the WB Bilateral remittance the reference year is 2020. For WB Bilateral remittance the reference year is 2018.

3.2 Empirical strategy

The empirical framework used in this study relies on a gravity model type approach. Gravity models are often used to explain flows between two countries, such as international trade flows.[18] More recently, gravity models have been used to explained FDI, migration and remittance flows.[19] In a gravity model, the flows are proportional to the country's economic sizes and inversely proportional to the distance between them. In our specific case, we replace the country's economic size with the size of the migrant population there (the migrant stock). We expect remittance flows between two countries to be proportional to the size of their migrant population and inversely proportional to the distance between them.

To understand the potential contribution of innovative data in explaining remittances flows we use the econometric specification described in Equation (4)

$$Remittance\ flow_{ij} = \beta_0 + \beta_1 UNDESA\ Stock_{ij} + \beta_2 \boldsymbol{Facebook\ Connections\ Indicator}_{ij} + \beta_3 GDP\ p.c.\ difference_{ij} + \beta_4 \boldsymbol{Facebook\ Connections\ Indicator}_{ij}\ X\ GDP\ difference_{ij} + \beta_5\ Distance_{ij} + \beta_6 Distance_{ij}^2 + \beta_7 Shared\ border + \varepsilon_{ij} \quad (4)$$

Where *Remittance flow*$_{ij}$ is the volume of remittances sent from country *i* to country *j* expressed in millions of dollars; *UNDESA Stock*$_{ij}$ is the stock of migrants from country j residing in country *i* while *GDP p.c. the difference* is the difference in GDP per capita between the two countries. These last two variables are used by the World Bank to define the weights used to produce their estimates and the Bilateral Remittance Matrix[20]. The *Facebook Connections Indicator*$_{ij}$ is expected to capture the social connections between countries beyond the conventional number of migrants residing in the country (a proxy of the stock of migrants residing in a country who actively maintain a connection with their place of origin). *Distance* and its squared term capture the distance between two countries. *Shared border* is a dummy variable that takes a value equal to one if two countries have a common border and zero otherwise. Finally, εij stands for unobserved effects. We standardised all nonbinary variables to achieve comparability between coefficients within and across specifications.[21]

In this empirical exercise, we rely on a cross-section dataset which, unlike panel data (multiple observations over time), does not allow us to control for origin–destination specific time-invariant effects which might affect

[18] The theoretical aspects of the gravity equation can be found in Anderson (1979), Deardorff (1984), Mátyás (1998) among others.

[19] See Portes and Rey (2005), Demekas at al. (2005) and Gupta and Mody (2006), Lueth and Riz-Arranz (2008), Kahouli and Maktouf (2015), Ahmed et al (2021) and Migali et al (2018) for recent applications of gravity model to equity, FDI, migration ad remittance flows.

[20] Ratha and Shaw (2007) state they use per capita GNI as a proxy for the migrant's income abroad. Here we use GDP which measures the goods and services produced within the country's geographical borders, by both country's residents and residents of the rest of the world. Instead, GNP measures the goods and services produced by only countries residents, both domestically and abroad..

[21] The normalised sample includes all observations entering at least in one of the four regressions.

both remittance flows and the set of independent variables included in the model. These and other potentially omitted variables might lead to biased estimates and the direction of the bias is hard to predict. Hence, the results presented in the next section should be interpreted with caution. We demonstrate the association between remittance flows and independent variables, not causality.

3.3 Regression results

We first estimate the basic form of the model presented in Equation (4), where the set of explanatory variables includes the stock of migrants, the difference in GDP per capita between the two countries on the two ends of remittance flow, a dummy variable equal to one if the two countries share a border, and distance between the countries with its squared term.[22] Here we exclude the Facebook Connections Indicator to obtain a set of results relying solely on the traditionally used set of variables. The dependent variable is the bilateral remittance flow and its elements, as described above. The regression results are presented in Table 3. Each column reports the results of a regression where the dependent variable is a different component of remittance flow (R1, R1WR, D752, and D752W in Columns 1, 2, 3 and 4 respectively) while the set of the explanatory variables is constant.

In all four cases, the migrant stock is positively associated with the size of remittance flow and the association is statistically significant. The larger the difference in GDP per capita between two countries, then the larger the remittance flow between them whereas the distance between countries is negatively (though not always statistically significantly) associated with remittance flows and its elements. This is in line with our expectations based on trade models (remittance flows between two countries are proportional to the size of the stock of migrants and inversely proportional to the distance between them). The coefficient of the difference in GDP per capita is positive and significant in all specifications, which indicates that larger differences in per capita income levels between countries (proxied by GDP per capita) are associated with larger remittance flows. The reported R squared suggests that this simple model explains about two-thirds of the variation in the main dependent variable of interest Personal Remittances reported in Column (1) and half of the variation for Personal transfers (D752) and Workers' remittances (D752W). The model has lower explanatory power in the case of Workers remittances ad compensation of employees (R1W1). Having a common border has a positive impact on Workers remittances and compensation of employees (R1W1), though negative for Personal transfers (D752) and Workers' remittances (D752W).

Table 3. Basic specification: Restricted set of explanatory variables (without the Facebook Connections Indicator)

	(1) R1	(2) R1WR	(3) D752	(4) D752W
Stock of migrants (1000)	0.864***	0.326***	0.686***	0.668***
	(0.027)	(0.017)	(0.014)	(0.014)
GDP per capita difference	0.082***	0.072***	0.053***	0.051***
	(0.025)	(0.017)	(0.014)	(0.014)
Distance	-0.046	-0.018	-0.045***	-0.043**
	(0.028)	(0.021)	(0.017)	(0.017)
Distance sq.	0.015	0.006	0.010	0.009
	(0.024)	(0.015)	(0.013)	(0.013)
Share border	-0.009	1.152***	-0.437***	-0.490***
	(0.126)	(0.095)	(0.077)	(0.077)
Constant	-0.017	-0.048**	0.010	0.002
	(0.038)	(0.023)	(0.019)	(0.019)
Observations	586	2,868	2,866	2,621
R-squared Adj.	0.67	0.20	0.46	0.48

Note: The dependent variable is the volume of remittance flow or its components (R1, R1WR, D752 and D752W) from country i to country j. The regressions are run for the sample of 2020. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

[22] One could run the regression using the log transformation of the dependent and independent variables. However, as task of the empirical exercise is to understand the potential of innovative data in explaining remittance flows, the estimation of elasticities remain beyond the scope of this study.

Then we introduce the variable on Facebook connections (Table 4). The coefficient at the Stock of migrants is both positive and statistically significant at the one percent level, though the magnitude is lower compared to those reported in Table 4. However, in all cases except for specification for Workers remittances and compensation of employees (Column 2) the coefficient at the Stock of migrants is larger in magnitude compared to the magnitude of coefficients at the Facebook Connections Indicator, meaning that for most of the specifications, Stock of migrants is the better predictor. As a result, the *Stock of migrants* variable is confirmed as the primary explanation of bilateral remittance flows. However, the Facebook Connections Indicator does contribute to explaining the bilateral remittance flow, although to a lesser extent than the stock of migrants. This extended specification has better explanatory power than the basic specification. The adjusted R-squared improved in all four specifications. For Personal remittances (R1) which is the subject of the main interest here, the model explains 78 percent of the variation in bilateral remittance flows.

We also include the interaction between the difference in GDP per capita and the Facebook Connections Indicator. We do so to hint at the direction of the Facebook connections *ij* which are by construction symmetric (the numerator is symmetric and does not recognise direction). For example, Facebook connections *ij* captures the number of connections between Albania and Italy but does not specify whether those are Albanians in Italy or Italians in Albania. We expect the sign of GDP per capita difference (positive or negative) to address this issue. Also, Distance between the countries preserves the pattern observed for the first set of regressions: the coefficients are negative though statistically significant for the specifications with dependent variable Personal transfers (D752) and Workers' remittances (D752W) only. Sharing a common border is positively associated with Personal remittances (R1) and Workers remittances and compensation of employees (R1WR) and negatively with Personal transfers (D752) and Workers' remittances here (D752W). One concern that arises in this specification is that the model might be affected by multicollnearity[23]. The multicollinearity test rejected the presence of multicollinearity in our model in all specifications.[24]

Table 4. Extended specification including Facebook Connections Indicator

	(1) R1	(2) R1WR	(3) D752	(4) D752W
Stock of migrants (1000)	0.516***	0.070**	0.506***	0.490***
	(0.031)	(0.028)	(0.023)	(0.022)
Facebook Connections Indicator (1000)	0.319***	0.167***	0.120***	0.114***
	(0.054)	(0.024)	(0.019)	(0.019)
GDP per capita difference	0.139***	0.109***	0.078***	0.075***
	(0.022)	(0.017)	(0.014)	(0.014)
Facebook Connections Indicator X GDP per capita difference	0.267***	0.284***	0.199***	0.200***
	(0.052)	(0.025)	(0.019)	(0.020)
Distance	-0.033	-0.018	-0.045***	-0.043**
	(0.023)	(0.020)	(0.017)	(0.017)
Distance sq.	0.015	0.007	0.011	0.010
	(0.019)	(0.015)	(0.012)	(0.013)
Share common border	0.170	1.258***	-0.374***	-0.413***
	(0.106)	(0.094)	(0.076)	(0.076)
Constant	-0.010	-0.059***	0.003	-0.008
	(0.032)	(0.022)	(0.018)	(0.019)
Observations	586	2,868	2,866	2,621
R-squared Adj.	0.78	0.24	0.48	0.50

Note: The dependent variable is the volume of remittance flow or its components (R1, R1WR, D752 and D752W) from country i to country j. The regressions are run for the sample of 2020. The variables are normalised. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.

[23] Two problems associated with multicollinearity are: (i) regression model estimates can become very sensitive to small changes in the set of other independent variables in the model and (ii) reduction in the precision of estimated coefficients.

[24] The Variance Inflator Factor (VIF) is below 2 in all specification which is much lower than the threshold suggested by the rule of thumb (10).

We also run the model by including two sets of the country of residence (origin of remittance flows) and country of origin (destination of remittance flows) dummies. These capture unobserved country-specific effects that might be correlated with the dependent variable and the independent variables, leading to biased estimates. The direction of the bias is hard to predict. For instance, in countries with developed digital infrastructure people are more likely to have internet access and social network accounts, resulting in more active Facebook accounts (Spyratos et al. 2019). In these countries, people may also be more likely to have an online bank account or a digital wallet, which impacts the statistics on official remittance flows and its elements because the transfer of financial resources is more traceable than informal money transfer methods.[25]

Table 5 reports the results. The coefficients at the Stock of migrants remain positive and statistically significant and are of similar magnitude. The coefficients at the Facebook Connections Indicator are also positive and statistically highly significant. For specifications on Personal Remittances (R1) and Workers remittances and compensation of employees (R1WR), the magnitude of the coefficients reported are smaller, but they are larger for Personal transfers (D752) and Workers' remittances (D752W) indicates that the omitted country-specific effects indeed were leading to a bias in estimates. Moreover, the direction of the bias varies depending on the measure of remittance flows. The coefficients reported for the interaction term of Facebook Connections Indicator with GDP per capita difference remain positive and statistically significant and even increase in magnitude. Instead, the coefficient at GDP per capita difference loses its significance which is perhaps because the country of origin and residence-specific dummies capture the effect in GDP per capita difference. Sharing a common border continues to be positively associated with Personal remittances (R1) and Workers remittances and compensation of employees (R1WR) and negatively with Personal transfers (D752) and Workers' remittances here (D752W)[26].

Table 5. Extended specification. Sending (residence) and receiving (origin) countries fixed effects included

	(1) R1	(2) R1WR	(3) D752	(4) D752W
Stock of migrants (1000)	0.493***	0.080***	0.490***	0.463***
	(0.032)	(0.029)	(0.024)	(0.024)
Facebook Connections Indicator (1000)	0.248***	0.158***	0.140***	0.146***
	(0.066)	(0.029)	(0.024)	(0.024)
GDP per capita difference	-0.223	0.156	-0.608	-0.352
	(29.189)	(46.539)	(38.415)	(44.618)
Facebook Connections Indicator X				
GDP per capita difference	0.244***	0.318***	0.213***	0.226***
	(0.064)	(0.027)	(0.021)	(0.022)
Distance	0.008	-0.093	-0.108	-0.198*
	(0.176)	(0.126)	(0.112)	(0.107)
Distance sq.	0.023	0.025	0.083	0.093*
	(0.089)	(0.062)	(0.057)	(0.053)
Share common border	0.247**	1.267***	-0.379***	-0.438***
	(0.112)	(0.102)	(0.084)	(0.084)
Constant	-0.156	0.147	0.279	0.516
	(0.219)	(0.780)	(0.651)	(0.757)
Country of residence FE	Yes	Yes	Yes	Yes
Country of origin FE	Yes	Yes	Yes	Yes
Observations	586	2,868	2,866	2,621
R-squared	0.82	0.24	0.47	0.49

Note: The dependent variable is the volume of remittance flow or its components (R1, R1WR, D752 and D752W) from country i to country j. The regressions are run for the sample of 2020. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.

[25] Unfortunately, our dataset is a cross-section (not a panel of repeated observation over several years) and we cannot include country pair specific dummies which would capture the time-invariant country pair fixed effects.

[26] For countries sharing a common border, it is more likely that people can become commuters or seasonal workers without residing there.

3.4 Robustness checks

The results presented so far indicate that innovative data have the potential to contribute to explaining bilateral remittances flows. The Facebook Connections Indicator and its interaction with GDP per capita difference are positive and statistically significant in all four specifications. To check the validity of these results we run a few robustness checks.

First, we restrict our sample to observations with positive remittance flows, hence, excluding all those observations for which the remittance (or its elements) flow is equal to zero. The results are reported in Table A1 (Annex 3). Here as well the estimates for the Stock of migrants are positive and statistically significant. The coefficients are of similar magnitude. Coefficients for Facebook Connection Indicators and its interaction with the GDP per capita difference - positive, statistically significant, and larger in magnitude. However, as in the main specification reported in Table 5, the GDP per capita difference turns insignificant in all four specifications. Sharing border is positively associated with Personal Remittances (R1) and Workers remittances and compensation of employees (R1WR) and negative with Personal transfers (D752) and Workers' remittances (D752W).

Second, in all four specifications (R1-D752W) we replace the dependent variables with the remittance flow estimates coming from the World Bank Bilateral Remittance Matrix. As 2018 is the latest year for which this dataset is available we run the regressions for 2018 using the same sample of country pairs for which the main specification was run. The objective is to see if the innovative data derived variables (Facebook Connections Indicator) can be used to explain the variation in remittance flow across countries using the principal dataset on remittance flows used by experts and the scientific community. The results are reported in Annex 3. The results show that also when replacing Eurostat reported figures with those coming from the World Bank Bilateral Remittance Matrix the coefficient at the Stock of migrants is positive and statistically significant. Given that the dependent variable is the same, the difference in the magnitude of the coefficients comes from differences in the sample. The fact that the stock of migrants is positive is not surprising as the bilateral flows are estimates using the Stock of migrants, as explained in Section 2. For all five subsamples, the Facebook Connection Indicator is positive and statistically significant, with a magnitude lower than the one observed for the Stock of migrants. The interaction term with GDP per capita is positive and even statistically significant for the whole sample reported in Column 6.

Overall the additional robustness checks confirm the validity of our results and support the hypothesis that data on connections between migrants and their countries and communities of origin can contribute to explaining bilateral remittance flows beyond the Stock of migrants and a set of other traditionally used explanatory variables. The Facebook Connections Indicator is positively and statistically significantly associated with remittance flows and its elements both using Eurostat data and WB Bilateral Remittances Matrix. Hence, innovative data can contribute to refining estimates of bilateral remittance flows.

4 Conclusions and recommendations

This report makes a novel contribution to understanding the factors shaping the volume of international bilateral remittance flows. Existing calculations of remittance flows are based on three main data sources: (a) current account and capital transfer data, (b) migrant stock data, and (c) socio-economic indicators. Our analysis shows that data on international social connections between migrants and their countries of origin are also positively associated with bilateral remittance flows. This suggests there is potential for also using data on connections between migrants and their countries of origin to refine existing estimates.

Our approach has been based on a statistical analysis which incorporates innovative data alongside data on migrant stocks and economic and geographic variables into an explanation of bilateral remittance flows. The data is Facebook's Social Connectedness Index, which provides a new opportunity to examine social connections on a global scale which has not been previously possible.

The analysis results confirm that data on the migrant population in a given country (the migrant stock) is the primary variable which plays a major role in explaining remittance flows from that country to a particular place of origin. Moreover, the results also show that the incorporation of data on migrants' international social connections leads to greater explanatory power of the statistical model. These findings are robust and statistically significant. This has been made possible by the availability of innovative data sources with a global coverage.

We recommend further exploration of the relationship between migrants' connections and remittance flows, innovative data sources and the contribution of both to informing future bilateral remittance flow estimations. In particular, this could include:

Global air traffic: the number of journeys made between countries can be an indicator of the connections between people living and working in each. Air traffic data can be accessed directly from private companies such as IATA[27] and Sabre.[28] Proxy data can also be accessed through Eurostat, within annual International Trade in Services Statistics.[29] The JRC has also produced an open-source dataset on mobility data for air traffic passenger data. This includes tourism and air traffic passenger data and is publicly available online.[30]

Mobile money-transfer services: mobile money transfer services are expanding rapidly, and could potentially provide a new data source on the volume and direction of flows between countries. There are already examples of this data being analysed in studies in France and Rwanda.

Google trends: geo-referenced online search data from Google Trends[31] can show the prevalence of certain keywords in different places. This could potentially be used to explore the extent to which people who reside in one country search for information about the situation in another. It has already been used to examine migration flows over time and to map diasporas (Böhme et al 2020; Draskovic 2020). Each search could also be a sign of an ongoing connection between a migrant and their country of origin, as they seek to keep informed about life there.

These exploratory activities would be a fruitful activity to pursue in the context of the International Working Group on Improving Remittance Flow Data.

[27] For more information see: https://www.iata.org/en/services/statistics/passenger-traffic-data/

[28] Sabre Corporation reports on a monthly basis the total number of passengers flying between countries (Sabre 2022). This data has been used to map trends in global human mobility, seasonal migration, and natural disaster-induced displacement (Recchi el al 2019; Gabrielli et al 2019; Rayer 2018). More recently it was used to study the relationship between the COVID-19 pandemic and international travel (Iacus et al 2020)

[29] Eurostat provided information on air transport (passengers) through International trade in services (bop_its6_det) is measured in euros. It has a limited geographic coverage (EU Member states and selected non EU countries) which limits its application for an empirical exercise at global scale. For more information, see: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=International_Trade_in_Services_statistics_-_background#Scope

[30] The dataset and metadata can be found in the JRC data catalogue: https://migration-demography-tools.jrc.ec.europa.eu/catalogue/dataset/ds00162, accessed 9th June 2022

[31] For more information see: https://trends.google.com/trends/

References

Ahmed, J., Mughal, M., and Martínez-Zarzoso, I. (2021). Sending money home: Transaction cost and remittances to developing countries. The World Economy, 44(8), 2433-2459.

Albayrak, M. B. K., Özcan, İ. Ç., Can, R., and Dobruszkes, F. (2020). The determinants of air passenger traffic at Turkish airports. Journal of Air Transport Management, 86, 101818.

Alvarez, Sandra Paola; Briod, Pascal; Ferrari, Olivier and Rieder, Ulrike (2015) Remittances: How reliable are the data? in Migration Policy Practice, V:2, pp. 42-46

Anderson, J. E. (1979). A theoretical foundation for the gravity equation. The American economic review, 69(1), 106-116.Deardorff, A. V. (1984). Testing trade theories and predicting trade flows. Handbook of international economics, 1, 467-517.

Avendaño, Rolando , Gaillard, Norbert and Nieto Parra, Sebastián (2009) Are Workers' Remittances Relevant for Credit Rating Agencies? (Paris: OECD Development Centre Working Papers)

Bailey, M., D. Johnston, T. Kuchler, D. Russel, B. State, and J. Stroebel (2020). The determinants of social connectedness in Europe. In S. Aref, K. Bontcheva, M. Braghieri, F. Dignum, F. Giannotti, F. Grisolia, and D. Pedreschi, (Eds.), Social Informatics, Cham, pp. 1--14. Springer International Publishing.

Bailey, M., R. Cao, T. Kuchler, J. Stroebel, and A. Wong (2018). Social Connectedness: Measurement, Determinants, and Effects. Journal of Economic Perspectives 32(3), 259—280

Beine, M., and C. Parsons (2015), "Climatic Factors as Determinants of International Migration," *Scandinavian Journal of Economics* 117(2): 723-767.

Beine, M., S. Bertoli, and J. Fernandez-Huertas Moraga (2016), "A Practitioners' Guide to Gravity Models of International Migration," *The World Economy* 39(4): 496-512.

Böhme, Marcus H; Gröger, André and Tobias Stöhr, (2020) Searching for a better life: Predicting international migration with online search keywords, in *Journal of Development Economics*, 142

Bosco, C., Grubanov-Boskovic, S., Iacus, S., Minora, U., Sermi, F. and Spyratos, S.,(2022) Data Innovation in Demography, Migration and Human Mobility, EUR 30907 EN, Publications Office of the European Union, Luxembourg, 2022, ISBN 978-92-76-46702-1, doi:10.2760/958409, JRC127369.

Choo, Y. Y. (2018). Immigration and inbound air travel demand in canada. Journal of Air Transport Management 71, 153--159.

Demekas, D. G., Horváth, B., Ribakova, E., and Wu, Y. (2005). Foreign Direct Investment in Southeastern Europe: How (and how much) can policies help?

Draskovic, Drasko (2020) Can Google search queries help countries map their diasporas? (UNDP, available online at https://www.undp.org/serbia/blog/can-google-search-queries-help-countries-map-their-diasporas, accessed 9th June 2022)

Eurostat (2021) Guide to statistics in European Commission development cooperation 2021 edition (Luxembourg: Publications Office of the European Union)

Fenoll, A. A., and Kuehn, Z. (2018). Immigrant networks and remittances: Cheaper together?. World Development, 111, 225-245.

Gabrielli, L., E. Deutschmann, F. Natale, E. Recchi, and M. Vespe (2019). Dissecting global air traffic data to discern different types and trends of transnational human mobility. EPJ Data Science 8(1), 26.

Greenwell, B. M. and Boehmke, B. C. (2020). Variable Importance Plots—An Introduction to the 'vip' Package. The R Journal, 12(1):343–366.

Gupta, P. (2005). Macroeconomic determinants of remittances: evidence from India.

Hagen-Zanker, J., and Siegel, M. (2007). The determinants of remittances: A review of the literature.

Iacus, S. M., Natale, F., Santamaria, C., Spyratos, S., and Vespe, M. (2020). Estimating and projecting air passenger traffic during the COVID-19 coronavirus outbreak and its socio-economic impact. Safety Science, 129, 104791.

IMF (2009a) International Transactions in Remittances: Guide for Compilers and Users (Washington DC: International Monetary Fund)

IMF (2009b) Balance of Payments and International Investment Position Manual Sixth Edition (Washington DC: International Monetary Fund)

Kahouli, B., and Maktouf, S. (2015). The determinants of FDI and the impact of the economic crisis on the implementation of RTAs: A static and dynamic gravity model. International Business Review, 24(3), 518-529.

KNOMAD (2017) Migration and Development Brief 28: Migration and Remittances (New York: KNOMAD and World Bank)

KNOMAD (2022) Remittance data working group (available online at https://www.knomad.org/remittance-data-working-group, accessed 1st June 2022)

Lueth, E., and Ruiz-Arranz, M. (2008). Determinants of bilateral remittance flows. The BE Journal of Macroeconomics, 8(1).

Lubambu, Karine Manyonga Kamuleta (2014) The Impact of Remittances on Developing Countries (Brussels: Directorate General for External Policies of the Union)

Mátyás, L. (1998). The gravity model: Some econometric considerations. World Economy, 21(3), 397-401.

Migali, S., Natale, F., Tintori, G., Kalantaryan, S., Grubanov-Boskovic, S., Scipioni, M., ... & Barbas, T. (2018). International migration drivers. Joint Research Centre: Ispra, Italy.

Obrzut, Robert (2020) Remittances according to the BPM6 manual. Eurostat https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Remittances_according_to_the_BPM6_manual

OECD (2006) International Migrant Remittances and their Role in Development, in International Migration Outlook (Paris: OECD)

Portes, R., & Rey, H. (2005). The determinants of cross-border equity flows. Journal of international Economics, 65(2), 269-296.

Ratha, D., & Shaw, W. (2007). South-South migration and remittances (No. 102). World Bank Publications.

Rayer, S. (2018). Estimating the migration of Puertoricans to Florida using flight passenger data. Bureau of Economic and Business Research. University of Florida.

Recchi, E., Deutschmann, E., and Vespe, M. (2019). Estimating transnational human mobility on a global scale. Robert Schuman Centre for Advanced Studies Research Paper No. RSCAS, 30.

Roberts, K. D., and Morris, M. D. (2003). Fortune, Risk, and Remittances: An Application of Option Theory to Participation in Village-Based Migration Networks 1. International Migration Review, 37(4), 1252-1281.

Rosales, R. (2020). Fruteros: Street vending, illegality, and ethnic community in Los Angeles. University of California Press.

Sabre (2022). Market intelligence, global demand data. https://www.sabre.com/products/market-intelligence/, Accessed: 19 May 2022

Schiopu, I. C., and Siegfried, N. (2006). Determinants of workers' remittances: Evidence from the European neighbouring region.

Spyratos, S., Vespe, M., Natale, F., Weber, I., Zagheni, E., and Rango, M. (2019). Quantifying international human mobility patterns using Facebook Network data. PloS one, 14(10), e0224134.

Spyratos, S., Vespe, M., Natale, F., Weber, I., Zagheni, E., and Rango, M. (2019). Quantifying international human mobility patterns using Facebook Network data. PloS one, 14(10), e0224134.

Tjaden, J. (2021, Dec). Measuring migration 2.0: a review of digital data sources. Comparative Migration Studies 9(1), 59.

UNDESA (2020) International Migrant Stock 2020 Documentation. Population Division POP/DB/MIG/Stock/Rev.2020 December. United Nations Department of Economic and Social Affairs, Population Division2020https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/undesa_pd_2020_international_migrant_stock_documentation.pdf

United Nations Department of Economic and Social Affairs, Population Division (2020). International Migrant Stock 2020. https://www.un.org/development/desa/pd/content/international-migrant-stock

Zagheni, Emilio (2020) How Facebook is helping to measure migration (Max Planck Gesellschaft available online at https://www.mpg.de/14931683/how-facebook-is-helping-to-measure-migration, accessed 1st June 2022)

Zien, A., Krämer, N., Sonnenburg, S., and Rätsch, G. (2009). The feature importance ranking measure. In Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pages 694–709. Springer

List of tables

Table 1. Tabular presentation of the definitions of remittances 6

Table 2. Summary statistics 11

Table 3. Basic specification: Restricted set of explanatory variables (without the Facebook Connections Indicator) 12

Table 4. Extended specification including Facebook Connections Indicator 13

Table 5. Extended specification. Sending (residence) and receiving (origin) countries fixed effects included ..14

Table A1. The subsample of observations with non zero values for the dependent variable.....23

Table A2. World Bank's Bilateral Remittance matrix, various subsamples 24

Annexes

Annex 1. Balance of payment items reported by Eurostat

Code	Label	Included in the study
SCAZ	Services: All modes of transport; Passenger; Of which: payable by border, seasonal, and other short-term workers	No
SDA1	Services: Travel; Business; Acquisition of goods and services by border, seasonal, and other short-term workers	No
D1	Primary income: Compensation of employees	No
D5Z	Secondary income: Current taxes on income, wealth, etc. payable by border, seasonal, and other short-term workers	No
D61Z	Secondary income: Social contributions payable by border, seasonal, and other short-term workers	No
D752	Secondary income: Personal transfers (Current transfers between resident and non resident households)	Yes
D752W	Secondary income: Workers' remittances	Yes
R1	Supplementary item: Personal Remittances	Yes
R1WR	Supplementary item: Workers remittances and compensation of employees	Yes

Annex 2. Data sources used in the study

Data	Source
Bilateral Remittance flows	World Bank
R1 Personal Remittances	Eurostat, Personal transfers and compensation of employees [bop_rem6]
R1WR Workers remittances and compensation of employees	Eurostat, Personal transfers and compensation of employees [bop_rem6]
D752 Secondary income: Personal transfers (Current transfers between resident and non-resident households)	Eurostat, Personal transfers and compensation of employees [bop_rem6]
D752W Secondary income: Workers' remittances	Eurostat, Personal transfers and compensation of employees [bop_rem6]
Facebook Connections Indicator	Own calculation based on Facebook's Social Connectedness Index
Stock of migrants	UNDESA, Population Division
GDP per capita, current prices	World Bank, development Indicators
Distance (between the most populated cities)	CEPII, GeoDist
Shared border	CEPII, GeoDist

Annex 3. Robustness checks

Table A1. The subsample of observations with non zero values for the dependent variable

	(1) R1	(2) R1WR	(3) D752	(4) D752W
Stock of migrants (1000)	0.547***	0.075*	0.490***	0.468***
	(0.045)	(0.039)	(0.035)	(0.036)
Facebook Connections Indicator (1000)	0.331***	0.181***	0.187***	0.190***
	(0.080)	(0.039)	(0.034)	(0.036)
GDP per capita difference	17.215	0.098	0.060	0.060
	(43.120)	(1.516)	(1.314)	(1.375)
Facebook Connections Indicator X GDP per capita difference	0.242**	0.335***	0.215***	0.230***
	(0.094)	(0.037)	(0.034)	(0.036)
Distance	-0.139	-0.179	-0.173	-0.242
	(0.253)	(0.175)	(0.174)	(0.184)
Distance sq.	0.034	0.026	0.105	0.111
	(0.129)	(0.081)	(0.087)	(0.086)
Share border	0.349**	1.106***	-0.341***	-0.420***
	(0.135)	(0.114)	(0.104)	(0.113)
Constant	2.075	0.091	0.508	0.689
	(5.694)	(1.462)	(1.284)	(1.345)
Country of residence FE	Yes	Yes	Yes	Yes
Country of origin FE	Yes	Yes	Yes	Yes
Observations	427	1,693	1,494	1,343
R-squared Adj.	0.83	0.21	0.47	0.47

Note: The regressions are run for the sample of 2020. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table A2. World Bank's Bilateral Remittance matrix, various subsamples

	(1) R1-subsample	(2) R1WR-subsample	(3) D752-subsample	(4) D752W-subsample	(5) (1)-(4)-subsample	(6) Whole sample
Stock of migrants (1000)	0.430***	0.377***	0.386***	0.375***	0.403***	1.000***
	(0.021)	(0.021)	(0.022)	(0.022)	(0.022)	(0.005)
Facebook Connections Indicator (1000)	0.286***	0.227***	0.246***	0.227***	0.237***	0.095***
	(0.047)	(0.021)	(0.022)	(0.022)	(0.021)	(0.005)
GDP per capita difference	3.924	0.048	0.061	0.054	0.051	0.027
	(5.022)	(0.554)	(0.569)	(0.696)	(0.566)	(0.093)
Facebook Connections Indicator X GDP per capita difference	-0.240***	-0.003	-0.000	-0.002	-0.007	0.167***
	(0.043)	(0.020)	(0.020)	(0.021)	(0.019)	(0.006)
Distance	-0.078	-0.336***	-0.457***	-0.365***	-0.364***	0.010**
	(0.124)	(0.092)	(0.102)	(0.102)	(0.093)	(0.005)
Distance sq.	0.013	0.086*	0.122**	0.083*	0.091**	-0.001
	(0.063)	(0.046)	(0.051)	(0.050)	(0.046)	(0.004)
Share border	0.292***	0.769***	0.741***	0.778***	0.719***	-0.403***
	(0.079)	(0.075)	(0.077)	(0.080)	(0.075)	(0.028)
Constant	-0.266*	0.716	1.038*	0.729	0.799	0.024
	(0.157)	(0.571)	(0.593)	(0.719)	(0.582)	(0.076)
Country of residence FE	Yes	Yes	Yes	Yes	Yes	Yes
Country of origin FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	586	2,868	2,866	2,621	2,937	23,490
R-squared Adj.	0.76	0.55	0.56	0.54	0.57	0.81

Note: The regressions are run for the sample of 2018. The variables are standardised. The dependent variable is always the bilateral flow from the Bilateral Remittance Matrix (World Bank). Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

- at the following standard number: +32 22999696, or

- by electronic mail via: https://europa.eu/european-union/contact_en

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications

You can download or order free and priced EU publications from EU Bookshop at: https://publications.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

KJ-NA-31118-EN-N

The European Commission's science and knowledge service

Joint Research Centre

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub

ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

EU Science, Research and Innovation

EU Science Hub

Publications Office of the European Union

doi:10.2760/318892

ISBN 978-92-76-53484-6