Article

Exploring the Macroeconomic Drivers of International Bilateral Remittance Flows: A Gravity-Model Approach

Giorgio Fagiolo *[ID] and Tommaso Rughi

Istituto di Economia, Scuola Superiore Sant'Anna, 56127 Pisa, Italy; tommaso.rughi@santannapisa.it
* Correspondence: giorgio.fagiolo@santannapisa.it

Abstract: This paper investigates the macroeconomic determinants of global bilateral remittance flows. Unlike existing studies, which have been often hampered by the lack of comprehensive and large-enough datasets, we use data originally covering 214 countries over the 2010–2017 period. We employ a gravity-model approach to explore the role played by dyadic and country-specific covariates in explaining remittances. We find that remittance flows are robustly and strongly impacted by size effects (i.e., number of migrants in the host country and population at home), transaction costs, common social, political, and cultural ties, output growth rate, and financial development at home. We also document the existence of a robust non-linear relationship between per capita income at home and remittance flows, both in the aggregate and across income groups.

Keywords: international remittances; international migration; gravity models

1. Introduction

Recent years have witnessed a spectacular increase in worldwide remittance flows[1]. As portrayed in Figure 1, panel (a), aggregate remittance flows have been steadily growing, in nominal terms, from USD 37 to around 550 billion (current) over the 1980–2019 period, which, in real terms, is equivalent to a six-fold increase. Furthermore, global remittances grew at a higher pace than merchandise trade, as documented by an almost doubled remittance/trade ratio. In the same period, the share of remittances to world GDP increased by 80%, as compared to a 55% rise in the trade-to-GDP ratio—see panel (b)[2]. Such an exponential surge in global remittances can only be partly explained by the increase in the number of (official) international migrants, which between 1980 and 2019 climbed from about 102 million to 270 million[3].

For many countries, especially low- and middle-income ones, remittance inflows have surpassed official development aid (ODA) as well as foreign direct investment (FDI), becoming their largest source of foreign exchange earnings. Overall, remittances make up a share of the country GDP ranging between 5 and 40 percent, which for some recipients is much larger than their export-to-GDP ratio[4].

Therefore, it comes as no surprise that remittances may have an extremely relevant impact on home-country economies, e.g., in alleviating poverty and contributing to development (Hagen-Zanker and Siegel 2007; Yang 2011). Remittances can indeed raise consumption and/or investment and play a crucial role in income smoothing, hence acting as an automatic stabilizer. However, they may also amplify the business cycle and thus destabilize economic activity (Cooray and Mallick 2013).

Since the seminal paper by Lucas and Stark (1985), several theoretical models have been laid out to explore the determinants underlying individual migrant remittance behaviors (Rapoport and Docquier 2006). A wide spectrum of micro-motives behind the reasons why migrants remit (and how much) have been investigated, ranging between the two extremes of pure altruism and self-interest, but also including tempered forms of altruism and strategic motivations (Carling 2008). These models have been taken to

the data mostly using two approaches. First, remittance determinants have been tested at the micro-level, focusing on single-country analyses and using household surveys[5]. Second, the macroeconomic drivers of aggregate remittances have been explored, employing panel data techniques, where the observational unit is a country and data come from balance-of-payments statistics across the years[6].
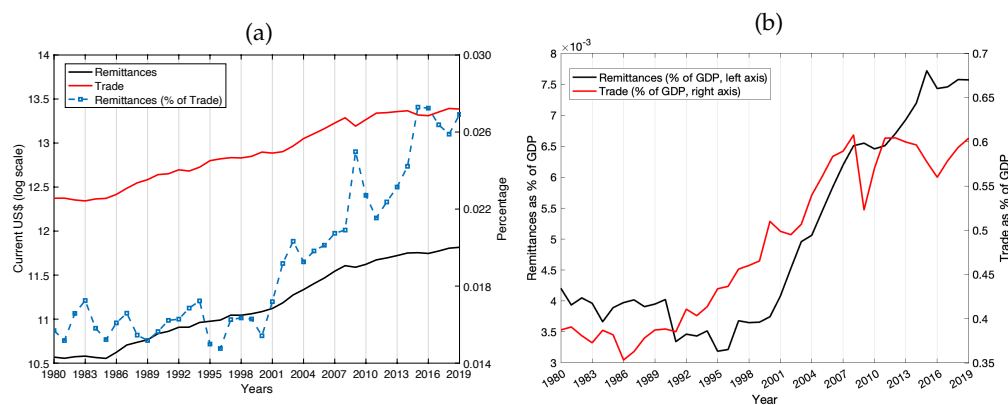

Figure 1. Aggregate flows of international remittances. (**a**): Aggregate world remittance and trade flows (current USD). (**b**): Remittances and trade as a percentage of world GDP. Source: Authors' calculation based on World Bank WDI data.

Another, albeit less common, stream of literature has exploited the inherent sending–receiving nature of aggregate remittance flows, fitting panel gravity models to bilateral (origin–destination) remittance flows[7]. A gravity-model approach to remittances has some value added as compared to panel-based econometric models. These include the possibility (i) to take into account separately home vs. host country-specific characteristics, therefore, proxying country masses and frictions, which may limit the volume of remittances due to transaction costs; (ii) to combine a microeconomic foundation with macroeconomic data, since a gravity-like relation emerges from simple micro-founded theoretical models (McCracken et al. 2017; Rapoport and Docquier 2006; Schiopu and Siegfried 2006). This last feature is particularly important, as sign predictions coming from the theory can be tested using real-world data.

However, existing gravity-based attempts to study international remittance flows have often been hampered by the lack of comprehensive and large-enough datasets. This has resulted in insufficient coverage for either the cross-sectional dimension (e.g., countries) or for the longitudinal one (e.g., years), or both.

In this paper, we attempt to overcome this limitation by fitting a panel gravity model to the "World Bank Migration and Remittance" database (Ratha and Shaw 2007). This repository originally contains estimates for international remittance flows from 214 sending (host) countries to 214 receiving (home) countries during the years 2010–2017. Unlike existing studies, the data we employ define a balanced "squared" panel, with all country pairs featured in each year, which may improve the robustness of the coefficient estimates. Obviously, some covariate may present missing values over different years, origins, and destinations, which may result in an unbalanced sample estimate. Nevertheless, we are able to fit our gravity model to a high number of remittance flows, covering a set of host and home countries which is generally larger than those employed so far in the literature (cf., Table A5 in the Appendix D)[8]. At the same time, the wide cross-sectional country coverage allows one to address the question of whether the estimated covariate elasticities differ among host/home country subgroups, e.g., if the determinants of remittance flows from rich or middle-income countries to poor ones are different from those underlying remittance flows in the whole sample.

Furthermore, we contribute to the literature employing gravity models to explain the macroeconomic determinants of bilateral international remittance flows in three additional

dimensions. First, unlike most of the existing papers[9], we explicitly deal with the zero-flow issue (Santos Silva and Tenreyro 2006), comparing results from OLS and Poisson pseudo-maximum likelihood (PPML) estimators. Second, we flexibly employ different sets of fixed effects (FEs), playing with alternative combinations of host/home and time FEs, so as to possibly mitigate omitted-variable biases. Finally, we test for non-linearities in the relation between income at home and remittance flows, to better explore the interplay between altruistic and self-interested motives.

Regression exercises, robustly across alternative specifications, document a strong size effect on remittances, as proxied by the number of migrants in the host country and population at home. Furthermore, remittances are impacted worldwide by bilateral frictions (i.e., transaction costs, common social, political, and cultural ties). Furthermore, we find that migrants at destination remit more the larger the growth rate and financial development at home. We also detect a robust non-linear relationship between per capita income at home and remittance flows, both in the aggregate and across income groups. In the whole sample, a U-shaped relationship emerges, suggesting that migrant behaviors seem to be driven by the desire to help relatives at home to deal with poverty and adverse shocks when per capita GDP at the destination is small (i.e., an altruistic motive, cf., Lucas and Stark 1985). Conversely, the association between income at home and remittances sent from middle-income/rich nations to poor countries follows an inverted U-shaped, in line with studies focusing on private transfers in less-developed countries (Cox et al. 1998).

The rest of the paper is organized as follows. Section 2 briefly discusses some theoretical background and existing papers dealing with a gravity-model perspective to the study of remittance flow determinants. Section 3 sets out our empirical strategy. In Section 4 we describe the data and methods employed. Section 5 presents our main results and reports on robustness checks. Finally, Section 6 presents our conclusions.

2. Theoretical Background and Related Literature

Disentangling the economic motivations behind migrants' decisions about if and how much to remit is not an easy task. Indeed remitting involves a large number of possible interacting determinants, having to do not only with the individual preferences and behavioral attitudes of the migrant, but also with economic, social, and political factors both at home and in the host country.

Since the seminal work by Lucas and Stark (1985), several theories have been proposed to fill the gap between two extreme views of migrant remittance behavior (see Yang 2011 and references therein). The first view considers remittances as driven by a purely altruistic motive, fueled by the migrant's desire to allow relatives back home to cope with poverty and adverse shocks. The second one models remittance behavior as stemming from a self-interested individual who only cares about their return to the community they left, and, therefore, remits to increase the likelihood to inherit and/or to buy assets at home. Migrants can however decide to remit because of reasons somewhat in between those two opposite motivations (Carling 2008; Hagen-Zanker and Siegel 2007), i.e., driven by a sort of tempered altruistic behavior, wherein migration provides mutual benefits for both the migrant and the family at home. In this framework, remitting can be the consequence of a sort of implicit contractual arrangement, whose motives include loan repayment (i.e., whenever migrants borrow money from their families to cover migration-related costs), exchange (e.g., compensation for child care provided to the migrant by recipients at home), and co-insurance (e.g., when negative shocks occur at home or when the migrant loses their job in the host country). Furthermore, the decision to remit can be induced by purely strategic motives Stark and Wang (2002), e.g., if skilled migrants have an incentive to send money back home to avoid further immigration of skilled workers, which might depress wages for skilled jobs.

Trying to empirically discriminate between these competing theories is not always possible. This is because, especially at the micro-level, alternative theories often predict similar signs as to the effect of covariates in econometric models explaining remittances.

In addition, poor data quality may prevent the design of an appropriate testing strategy (Rapoport and Docquier 2006). In fact, at the macro-level, the lack of high-quality data has been the major hurdle faced by researchers attempting to assess the relative importance of aggregate determinants of country remittance flows.

This is true in particular in the case of the dependent variable not being aggregate country (sent or received) remittances, but aiming at explaining bilateral international remittance flows between pairs of countries using a gravity-model approach. Such a framework is particularly appealing if the researcher aims at assessing the macroeconomic determinants of remittances. Indeed, remittance flows at the macro-level have an intrinsic sender–receiver essence. Therefore, they naturally lend themselves to a modeling setup where flows are explained using separate origin and destination characteristics, as well as features related to the dyadic interaction between host and home country, capturing the role of frictions induced by transaction costs. More importantly, a gravity specification for remittance flows emerges as the equilibrium prediction of a two-period model where migrants care about consumption and invest in a host country safe asset as well as in a home country risky asset (McCracken et al. 2017; Rapoport and Docquier 2006; Schiopu and Siegfried 2006). This allows one to derive precise implications about the expected sign of remittance macroeconomic drivers, stemming from microeconomic assumptions about altruistic vs. self-interested migrant behaviors. For example, the model predicts that, if migrants are sufficiently altruistic[10], remittances to relatives in the home country should decrease the larger the income at the destination (net of that in the origin).

Existing empirical evidence stemming from gravity-model exercises about the role of macroeconomic determinants of remittances is nevertheless not conclusive (cf., Table A5 in the Appendix D). Indeed, the availability of good-quality datasets featuring, for a large set of country pairs and years, all bilateral remittance flows has always been extremely poor, hence limiting the scope of applied analyses in this field. As we show in the Appendix D (Table A5), existing works using remittance gravity models (RGMs)[11] have usually focused on a limited number of sending and receiving countries (in the range of 16–89 and 7–75) observed for a short number of years, which is typically inversely related to the country sample size used in the analysis. Furthermore, the panel structure is often strongly unbalanced: the set of sending countries never coincides with that of the receiving ones, implying a rectangular dataset. This implies that one may not correctly evaluate the impact of country-specific determinants in the two-way remittance relationship between any two countries in the sample (we shall come back to this issue in Section 5.3).

Furthermore, sign predictions are indeed often contrasting and sometimes uncertain. For instance, the impact of economic conditions at home (i.e., country income, GDP growth, etc.), as well as that of transaction costs (as modeled using geographical distance and traditional gravity dyadic relations), may be biased and highly sensitive to the FE specification, the treatment of zero flows, and the presence of endogeneity. Additionally, the sample of countries included in the analysis—either as sending or receiving—greatly varies across exiting studies. Even in Ahmed et al. (2021), who employ the "World Bank Migration and Remittance" database as in the present work, only the most important migration corridors are considered[12].

In addition to heterogeneity in country sample size and composition, there are further issues limiting the comparability of results (and their robustness) across existing RGM works. First, all studies but Docquier et al. (2012) apply an OLS estimator only, thus excluding ex-ante the possibility of dealing with zero-remittance flows. It is well known that, under heteroskedasticity, this may imply biased estimates of the true elasticities (Santos Silva and Tenreyro 2006). Second, existing papers employ different assumptions as to the set of sending–receiving countries and time FEs[13]. As discussed at length in Anderson and van Wincoop (2003), failing to properly control for cross-sectional and time-varying origin–destination heterogeneity may lead to strong unobserved-variable biases.

3. Empirical Strategy

In our exercises, we instead explicitly address both the zero-flow and cross-sectional/time-varying origin–destination heterogeneity issues. Firstly, we check the robustness of our OLS estimates against those obtained via a PPML estimator, which explicitly includes zero flows in the estimation. Secondly, we experiment with different assumptions as to FEs employed in the regressions (more on that in Section 4). More specifically, we employ both a structural specification (where only bilateral country variables are included together with origin-time and destination-time FEs) and a less stringent one, where variation in the host country is controlled by origin-time FEs, while at home we employ time-invariant destination FEs and a battery of time-varying observable characteristics.

Furthermore, in order to tackle the small sample size and composition issues, here, we use data from the "World Bank Migration and Remittance" database, originally reporting estimates for international remittance bilateral flows from 214 sending (host) countries to 214 receiving (home) countries for the years 2010–2017. As discussed in more details below, we are eventually able to retain 176 countries in our regression analyses, after removing those that never remit nor receive and for which some covariates are missing. Therefore, we still cover most remittance flows in the world in a squared panel, i.e., all incoming and outgoing (zero or positive) remittance flows are present in the dataset in each year (see Table A3 in the Appendix C and cf., the discussion in Section 5.3). Despite World Bank data not being empirically observed but coming from estimates, this is, to the best of our knowledge, the best choice if one aims at a large country coverage for a sufficiently long number of years—we shall come back to this point in Section 6. It is worth noting that our enlarged country sample size allows us to run separate gravity regressions, where sending vs. receiving countries belong to subgroups defined according to their income (i.e., high, middle and low income), and, therefore, to assess how the aggregate drivers of remittances change depending on the development levels of the home and host countries. Finally, we run a set of robustness checks to explore for potential sources of bias in our estimates.

As mentioned, we aim at reassessing in a more robust way the role of macroeconomic drivers of international remittance flows. In our RGM approach, we control for three types of covariates, net of various combinations of origin country (i), destination country (j), and year (t) FEs. The first of these is the stock of migrants in the host country, which varies across origins and destinations of remittance flows (i.e., respectively, migrant host and home countries) and time (i.e., over the ijt triplet). In line with the existing literature (Freund and Spatafora 2008), we expect remittances to increase with the stock of migrants in the host country ("number of migrants" hereafter), due to a sheer size effect[14].

The second family of covariates are time invariant and vary across pairs of countries (i.e., across the ij dyad). These include geographical distance, contiguity, and typical gravity-model bilateral dummies capturing ties between home and host countries (i.e., common language, and religion, as well as the existence of any former colonial relationship). Remittances are expected to decrease with distance and increase if the host and home countries share a border, as they both proxy transaction costs[15]. The impact of contiguity may, however, be negative if, net of geographical distance, sharing a border enhances informal remittances and discourages formal ones, as travel costs are lower and migrants find it easier to remit by unofficially transferring money across borders (Lueth and Ruiz-Arranz 2006). We also expect holding ties with the host country to boost remittances. Indeed, migrants already speaking the host country's language or sharing the same religion may be more integrated in the new society and, hence, they may more easily get a job. Similarly, a common past of colonial relationships typically implies some degree of institutional similarity and political ties between home and host countries. This may facilitate migration towards the former colonizer and subsequent integration.

The third class of potential remittance determinants includes origin and destination country-specific factors, which vary both across countries and time (along the it and/or jt dimensions). More specifically, we focus on covariates proxying for country economic conditions (i.e., per capita GDP and GDP growth), size effects (i.e., population), agriculture

(i.e., share of rural population), education (i.e., expenditure share of education over GDP and enrollment rate), and efficiency of financial institutions (as proxied by the share of bank branches)[16]. Net of host country and other destination covariates, we expect remittances to increase (i) the larger the population size at home (as, net of the number of migrants at the origin, the greater will be the basin of potential recipients)[17]; (ii) the larger the share of home rural population and education level, as this may reflect loan repayment or exchange motives, and more generally that remittances are used for investment purposes rather than to boost consumption; cf., see Hagen-Zanker and Siegel (2007); Yang (2011)[18]; (iii) the more developed the financial system is at the destination, because this eases formal money transfers both at home and in the host country (Giuliano and Ruiz-Arranz 2009); and (iv) the lower the home GDP growth rate, as it may be correlated with a less dynamic economic environment at home and, therefore, may proxy for sender–receiver differences in the business cycles (Kakhkharov et al. 2017).

The impact of per capita GDP (as a proxy for income) on remittances is less straight-forward. Given the host country income, we expect that, if altruistic motives dominate, then a larger income at home would decrease remittance flows. Instead, if migrants are more self-interested and care about investment, an increasing income at home should boost remittances. However, as shown in Cox et al. (1998), altruistic and self-interested motives may switch as home income increases. This means that one may possibly observe U-shaped (or inversely U-shaped) relations when investigating the impact of income on remittances. For example, migrants from poor home countries may mostly remit to help relatives back home coping with poverty and adverse shocks. On the contrary, migrants from high-income countries may start remitting pushed by tempered altruistic or even purely self-interested motivations (e.g., loan repayment, exchange, or co-insurance). Therefore, in order to explore whether this is the case in our data, we test for possible non-linearities in the (home) income–remittance relationship.

A number of other macroeconomic determinants may potentially affect international remittance flows (Carling 2008). These include, among others, climate and disasters, interest and exchange rate differentials, and poverty and fragility indicators. However, as discussed in the next section (see also in Appendix B, Table A2), their detected impact in our regression exercises was neither conclusive nor robust across alternative specifications and estimation methods, and, therefore, they were discarded from the analysis.

4. Data and Methods

We fit to the data a panel gravity model whose non-linear baseline formulation reads:

$$R_{ij}^t = \kappa \exp\left\{\eta_i^t + \psi_j + \delta^t + \gamma D_{ij} + \phi M_{ij}^t + \theta X_j^t\right\}\epsilon_{ij}^t, \tag{1}$$

where R_{ij}^t are remittances (in levels) from i (origin/host country) to j (destination/home country) in year t; κ is a constant; η_i^t are origin-time FEs, ψ_j are destination, time-invariant FEs, and δ^t are time dummies; D_{ij} is a set of time-invariant, bilateral covariates; M_{ij}^t is the number of migrants, i.e., the stock of people born in country i and living in country j in year t; X_j^t is a battery of destination-specific, time-dependent regressors; and ϵ_{ij}^t are the errors.

In this baseline specification, we control for both cross-sectional and longitudinal variation across origins of the remittance flows, while focusing on time-varying observable characteristics of destination/home countries—once their unobservable cross-sectional differences are controlled for. As a consequence, country-specific variables only depend on destinations X_j^t and covariates that are specific to the origin and vary over time cannot be inserted, as they would be collinear to η_i^t. This means that their coefficient estimates are to be interpreted as "net of the country-origin covariates".

To check the robustness of the bilateral-variable coefficient estimates (especially as far as M_{ij}^t is concerned), we also fit a structural gravity specification reading:

$$R_{ij}^t = \kappa \exp\{\eta_i^t + \psi_j^t + \delta^t + \boldsymbol{\gamma} \boldsymbol{D}_{ij} + \boldsymbol{\phi} M_{ij}^t\}\epsilon_{ij}^t, \tag{2}$$

where η_i^t and ψ_j^t are, respectively, origin-time and destination-time fixed effects, and country-specific covariates $\boldsymbol{X}_j^t$ are omitted (more on that in Section 5.3).

Remittance and migrant data come from the "World Bank Migration and Remittance" database (cf., Appendix A, Table A1 for descriptions and sources of all variables used in our analysis). Bilateral matrices originally report estimated remittance flows (in millions of US dollars) from 214 sending (host) countries to 214 receiving (home) countries for the years 2010–2017—we shall go back to discussing some possible issues related to using estimated rather than observed data in Section 6.[19]. A possible concern is related to the non-stationarity of the dependent variable. To double check that this is not the case, we ran Harris–Tzavalis (HT) panel data unit root tests on remittance flows (Harris and Tzavalis 1999), after creating a unique identifier for origin–destination flows and making the resulting panel strongly balanced by dropping some observations. The HT test is in our case more appropriate than other alternatives (see, e.g., the Levin–Lin–Chu test) as it assumes that the number of panels is large while the number of time periods is fixed and small. The results indicate that the null hypothesis of non-stationary remittance flows is rejected[20].

We also employ data about migration stocks, which contain estimates of the number of people M_{ij}^t born in the destination country j and living in the host country i (i.e., the origin of the remittance flow from i to j) in the years 2010, 2013, and 2017. The estimates are based on the "Migration and Remittances Factbook" (various years) and are used here as a covariate controlling for bilateral migration-size effects at the origin. Since we do not have bilateral migration observations for all the years covered in the remittance database, we employ two alternative strategies. First, we estimate Equation (1) only for the three waves where both remittance and migration are available. In this setup, the number of migrants M_{ij}^t enters as a contemporaneous covariate for R_{ij}^t, $t = 2010, 2013, 2017$. We label this case in our results as "Year = t". Second, we fit our model in all the years for which we do have remittance data ($t = 2010, \ldots, 2017$), building a stepwise migrants-at-destination variable reading $\tilde{M}_{ij}^t = M_{ij}^{2010}$ for $t = 2010, 2011, 2012$, $\tilde{M}_{ij}^t = M_{ij}^{2013}$ for $t = 2013, \ldots, 2016$, and $\tilde{M}_{ij}^t = M_{ij}^{2017}$ for $t = 2017$. We label this case in our results as "stepwise" (see, however, Section 5.3 for a robustness exercise where we use linearly interpolated migrant stocks). Descriptive statistics for bilateral remittances and the number of migrants in three selected years (2010, 2013, and 2017) are reported in the Appendix C, cf., Table A4.

In addition to M_{ij}^t, we account for two sources of variation. The first one ($\boldsymbol{D}_{ij}$) includes usual bilateral, time-invariant, standard gravity regressors such as geographical distance, contiguity, common language, common religion, and the existence of any colonial relationship in the past. We have also experimented with additional bilateral, time-invariant effects such as common ethnic language, common currency, weighted versions of geographical distance, as well as different definitions of colonial relationships; see Appendix B, Table A2 for details. However, these covariates have been excluded from all tested specifications as they turned out to be not significant in almost all of our regressions and contained too many missing values for our selected sample of 176 countries.

The second one ($\boldsymbol{X}_j^t$ in Equation (1)) controls for destination country-specific factors that may affect remittance flows and vary both longitudinally and cross-sectionally. These are: income (as proxied by per capita GDP, pcGDP henceforth), GDP growth rates, population, the share of rural population, expenditure on education (as a % of GDP), enrollment rate, and the share of bank branches[21]. Since, as mentioned, we are interested in exploring possible non-linearities in the relation between home income and remittances, we also insert among our covariates the square of per capita GDP.

After removing countries that did not ever remit or receive in at least one year (12 in total), and those for which our selected covariates are seldom observed (26 countries), we end up with a final sample covering 176 countries (see Table A3 in the Appendix C) for the period 2010–2017. The remittance flow panel has a squared format, i.e., all in/out (zero and positive) remittance flows are featured in the dataset in each year (more on this point in Section 5.3).

We begin fitting Equation (1) with a standard OLS estimator. This requires us to log-linearize the gravity model and, therefore, does not allow one to account for zero-remittance flows that, as Table A4 in the Appendix C suggests, is a sensible issue in our data. It is well-known that under heteroskedasticity this implies potentially biased coefficient estimates (Santos Silva and Tenreyro 2006). Therefore, we check the robustness of our OLS baseline results estimating (1) with a Poisson pseudo-maximum likelihood (PPML) estimator, using remittances in level and including their zero observations.

5. Results

5.1. Whole-Sample Regressions

Table 1 reports coefficient estimates obtained when Equation (1) is fitted to whole-sample data. We show two sets of specifications. The first set—columns (1)–(4)—includes baseline OLS and PPML estimates where empirically observed values for both the number of migrants and per capita GDP are employed. As discussed in Section 4, we also experiment with two alternative setups as far as the number of migrants covariate is concerned, according to whether only years 2010, 2013, and 2017 are considered ("Year = t", cf., columns 1 and 2) or the "stepwise" version of M_{ij}^t is employed (columns 2 and 4).

Table 1. Regression results. Dependent variable: bilateral remittance flows. Whole-sample estimates of gravity-model coefficients (Equation (1)). Fixed effects specification: $\eta_i^t + \psi_j$. Columns (1–2): OLS estimates. Columns (3–4): PPML estimates. Columns (1, 3): the covariate "number of migrants" is observed only in years 2010, 2013, 2017. Columns (2, 4): The "stepwise" version of the covariate "number of migrants" is employed (see Section 4 for more details). Standard errors in round parentheses. Significance levels: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

	OLS		PPML	
	(1) Migrants at Destination Year = t	**(2)** Migrants at Destination Stepwise	**(3)** Migrants at Destination Year = t	**(4)** Migrants at Destination Stepwise
Migrants at Destination	0.993 *** (0.008)	0.992 *** (0.992) ***	0.858 *** (0.021)	0.837 *** (0.013)
Distance	−0.096 *** (0.019)	−0.104 *** (0.010)	−0.181 *** (0.040)	−0.243 *** (0.025)
Contiguity	−0.360 *** (0.069)	−0.355 *** (0.035)	−0.216 * (0.091)	−0.222 *** (0.053)
Common Language	0.040 (0.039)	0.049 * (0.021)	0.183 * (0.076)	0.207 *** (0.045)
Colonial Relationship	0.164 * (0.072)	0.002 (0.039)	0.093 (0.112)	0.006 (0.062)
Common Religion	0.101 * (0.048)	0.217 *** (0.025)	0.233 * (0.101)	0.331 *** (0.060)
pc GDP	−1.599 *** (0.186)	−1.420 *** (0.101)	−1.669 *** (0.448)	−2.479 *** (0.280)
pc GDP Squared	0.090 *** (0.010)	0.084 *** (0.005)	0.080 *** (0.024)	0.130 *** (0.015)
GDP Growth	−0.044 *** (0.005)	−0.021 *** (0.002)	−0.021 * (0.009)	−0.002 * (0.006)
Population	0.118 *** (0.009)	0.100 *** (0.005)	0.171 *** (0.019)	0.178 *** (0.012)
Rural Population Share	0.005 *** (0.001)	0.007 *** (0.001)	0.006 * (0.003)	0.008 *** (0.002)
Exp on Education (% of GDP)	0.045 *** (0.009)	0.075 *** (0.004)	0.015 (0.028)	0.033 * (0.016)
Enrollment Rate	0.013 *** (0.002)	0.007 *** (0.001)	0.018 *** (0.003)	0.017 *** (0.002)
Bank Branches	0.014 *** (0.001)	0.019 *** (0.000)	0.010 *** (0.001)	0.011 *** (0.001)
Obs	12,510	33,408	55,406	147,740
R^2	0.881	0.902	0.941	0.936
Prob > F	0.000	0.000	-	-

As a first general observation, both diagnostics and the signs of the estimated coefficients turn out to be very stable across our first four baseline specifications, i.e., the R^2 is always very high and we do not detect sign inconsistencies as the estimation method and the definition of the number of migrants covariate change.

Notice also that the number of observations actually fitted substantially varies across specifications. This is due to two related issues. First, the squared remittance matrix contains $176 \times 175 = 30{,}800$ observations per year, that is 246,400 observations in total, of which only 55,731 are strictly positive (cf., Appendix C, Table A4). When using the "Year = t" version for the number of migrants, there are only about 21,000 strictly positive observations notionally available with the OLS estimator. Second, the presence of missing values in the covariates (on average, about 40% of the observations), scattered across years and countries, further limits the number of observations actually available and makes the panel unbalanced. This implies that in the OLS case one can fit only about 12,500 observations in the "Year = t" case and about 33,400 observations when using the "stepwise" option. Notwithstanding missing values in the covariates, the sample size becomes much larger when we employ the PPML estimator, as all zero flows are considered. We will explicitly address the robustness of our results to the presence of missing values in the covariates in Section 5.3.

As far as the significance and magnitude of the coefficients are concerned, there appears to be a small subset of covariates that, overall, seem to impact remittances in a less robust way across specifications. For instance, colonial ties sometimes become not significant, whereas the magnitude of coefficient estimates for geographical distance, GDP growth, and common language occasionally change.

Nevertheless, the results in Table 1 suggest quite a robust and consistent pattern as to the association between macroeconomic determinants and bilateral remittances. Indeed, remittances increase the larger the pool of migrants at the origin, whether home/host countries share a language, colonial or religion tie, the larger the total population, rural population share, expenditure on education, enrollment rate, and bank branches at home. On the other hand, remittances decrease the larger the geographical distance between origin and destination and GDP growth at home, and when home and host countries share a border.

Taken together, the foregoing results imply a number of considerations as to the role played by alternative macroeconomic drivers of remittances. First, the stock of migrants at the origin appears to exert a very stable and strong size effect on remittance flows, net of the magnitude of the basin of recipients at home, controlled for by total population.

Second, the negative impact of geographical distance hints at transaction costs as being a relevant factor in explaining remittances. Note that sharing a border reduces formal flows of money towards home, which may indicate that contiguity could be an incentive to boost informal ways to remit (we shall further comment on this interpretation in Section 5.3). However, this interpretation holds only if World Bank estimates do not include informal remittances, which should not be taken for granted. Third, our exercises confirm that remittances are facilitated if the origin and destination countries hold social, cultural, and political ties, as this may further decrease transaction costs. Fourth, the positive impact of rural population share, expenditure on education, and enrollment rate suggests that, as mentioned in Section 2, remittances are employed relatively more for investment motives rather than as a way to boost consumption at home. Finally, relatively to the origin, a less dynamic but more financially developed home economy is able to attract more remittances.

We also detect a consistent and robust non-linear impact of income (pcGDP) at destination. More precisely, as Table 1 shows, we find that the relation between home income and remittances is U-shaped, with remittance flows decreasing for low-income levels and increasing for high-income ones. The U-shaped relationship between (per capita) income and remittances observed in the whole data sample, is depicted in Figure 2. There, we plot the marginal effect of pcGDP (across its observed range) on bilateral remittance flows on a

log–log scale (for the specification in Table 1, column 2) and we add in the background a histogram of the observed whole-sample distribution of pcGDP at destination.
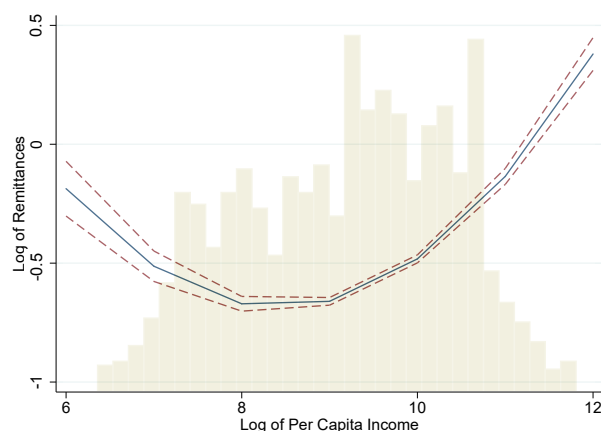
Figure 2. Non-linear marginal impact of (per capita) income (pcGDP) on remittances. Solid blue line: Whole sample OLS estimates from column (2), Table 1. Dashed red lines: 95% confidence bands. X-axis: log of (per capita) income (pcGDP) in the observed whole-sample range of the covariate. Y-axis: log of remittance flows. The histogram in the background depicts the whole-sample distribution of pcGDP at the destination across countries and years (bar heights are proportional to observed frequencies).

This evidence suggests that for relatively poor destination countries, altruistic motives dominate in the sending behavior of migrants (net of host income), whereas self-interest seems to be more relevant in the decision to remit as income at home—relative to that at the origin—becomes larger than a given threshold. This is in contrast with results obtained by Cox et al. (1998), who found that, in the case of private transfers in Peru, exchange motives prevail for low-income recipients and altruistic motives predominate for those with high incomes.

In order to dig further on this point—and more generally to better understand if the whole-sample results still hold when we consider subsamples of remittance flows—we now move to a more disaggregated analysis, where both the origin and destination countries are classified according to their income group.

5.2. Remittance Flows by Origin and Destination Income Group

We categorize countries in our sample into three income groups (poor, middle, and rich), using the 2020 WB income group classification based on the Atlas method[22]. In our exercises, countries are defined as (i) "poor" (PC) if they belong to the "low" or "lower-middle" WB income group, (ii) "middle" (MC) if they are classified as "upper-middle", or (iii) "rich" (RC) if they belong to the "high" WB income group.

This allows us to form nine non-overlapping subsamples for our dependent variable, according to whether the origin and the destination country of remittance flows are classified as PC, MC, or RC.

We are particularly interested in focusing on two subsamples, namely, those where remittances are sent to a PC either from an RC ("rich to poor") or an MC one ("middle to poor"). This is because of two main reasons. First, poor countries are those where remittances can impact the most in terms of alleviating poverty and promoting economic growth. Second, as discussed above, there exists literature (see, e.g., Cox et al. 1998) showing that in poor (home) countries exchange and self-interested motives prevail for low-income levels, while altruistic motives predominate when income grows. This suggests that an inverse U-shaped relationship might be observed when focusing on poor home countries, contrary to what we have found in our whole-sample estimates.

Table 2 summarizes our main outcomes. We report the OLS coefficient estimates and significance levels obtained when fitting Equation (1) to the two subsamples of interest, and comparing them with the whole-sample results from Table 1, columns (1)–(2)[23].

Table 2. Regression results by income group. Dependent variable: bilateral remittance flows. OLS estimates of gravity-model coefficients (Equation (1)) in subsamples defined according to the income group of the origin and destination remittance flow countries. Column (1): rich to poor; column (2): middle to poor; column (3): whole-sample estimates from columns (1)–(2) in Table 1. Fixed effects specification: $\eta_i^t + \psi_j$. Number of migrants: "Year = t" means that the covariate is observed only in years 2010, 2013, 2017; "stepwise" means that the stepwise version of the covariate is employed (see Section 4). Significance levels: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Dependent Variable: Bilateral Remittance Flows	Number of Migrants	Income Group (Origin → Destination)		
		(1)	(2)	(3)
		Rich → Poor	Middle → Poor	Whole Sample
Number of Migrants	Year = t	1.044 ***	1.035 ***	0.993 ***
	Stepwise	1.026 ***	1.037 ***	0.992 ***
Distance	Year = t	0.107 *	0.043 *	−0.096 ***
	Stepwise	0.117 ***	0.029 *	−0.104 ***
Contiguity	Year = t	−0.029	−0.388 *	−0.360 ***
	Stepwise	−0.122	−0.406	−0.355 ***
Common Language	Year = t	0.146	0.279 *	0.040 *
	Stepwise	0.075	0.240 ***	0.049 *
Colonial Relationship	Year = t	0.287	0.074	0.164 *
	Stepwise	0.210 *	0.004	0.002
Common Religion	Year = t	−0.593	−0.122	0.101 *
	Stepwise	−0.088	0.161 *	0.217 ***
pcGDP	Year = t	1.709 *	2.146 *	−1.599 ***
	Stepwise	1.644 ***	1.158 *	−1.420 ***
pcGDP Squared	Year = t	−0.132 **	−0.154 *	0.090 ***
	Stepwise	−0.109***	−0.073 *	0.084 ***
GDP Growth	Year = t	−0.140 ***	−0.152***	−0.044 ***
	Stepwise	−0.096 ***	−0.097 ***	−0.021 ***
Population	Year = t	0.282 ***	0.326 ***	0.118 ***
	Stepwise	0.232 ***	0.255 ***	0.100 ***
Rural Population Share	Year = t	0.004 *	0.001 *	0.005 ***
	Stepwise	0.004 ***	0.005 **	0.007 ***
Exp on Education (% of GDP)	Year = t	0.034 *	0.058 **	0.045 ***
	Stepwise	0.042 ***	0.039 ***	0.075 ***
Enrollment Rate	Year = t	0.015 ***	0.012 **	0.013 ***
	Stepwise	0.009 ***	0.003 *	0.007 ***
Bank Branches	Year = t	0.048 ***	0.051 ***	0.014 ***
	Stepwise	0.039 ***	0.039 ***	0.019 ***
Obs	Year = t	2293	1009	12510
	Stepwise	5561	3604	33,408
R^2	Year = t	0.937	0.938	0.881
	Stepwise	0.990	0.983	0.941
Prob > F	Year = t	0.000	0.000	0.000
	Stepwise	0.000	0.000	0.000

To begin with, note that the sign and significance of most of the macroeconomic drivers of remittances, as identified in the whole data sample, are confirmed also when disaggregating the data by origin and destination income group. In particular, size effects exerted by migrants in the host country and home-country population continue to be strong determinants of remittances also in the "rich to poor" and "middle to poor" subsamples.

The same observation applies for GDP growth, rural population share, education-related covariates, and bank branches.

On the contrary, the impact on remittances of macroeconomic drivers associated with transaction costs and bilateral ties differs between whole-sample (column 3) and income group (columns 1–2) regressions. On the one hand, a larger geographical distance now boosts remittances from rich to poor countries. This is in line with the evidence found in de Sousa and Duval (2010) for Romania, and is in general consistent with the idea that migrants from poor countries, who travel longer distances, remit more because they cannot visit their home country very frequently—and, therefore, they cannot carry in-kind or cash with them. In addition, distance may be positively associated with remittances to poor countries due to a loan repayment motive: if family members living in distant, low-income countries partly covered the higher migration cost, such a loan may be repaid in the form of larger remittances thereafter. These interpretations are also consistent with the almost not significant effect of contiguity on remittances from high to middle income countries to poor ones, also because poor countries do not tend to share a border with richer ones (i.e., informal remitting channels become irrelevant). Nevertheless, our disaggregated regressions do not pick up any positive association between contiguity and formal remittance flows in income-disaggregated samples, as perhaps one may have expected.

Second, common ties related to religion, former colonial relationships, and (partly) language become much less important than in the aggregate while explaining remittances from middle and rich countries to poor ones. Although the interpretation of why this happens is less straightforward, the significance loss in social, cultural, and political ties as macroeconomic drivers of remittances may be in line with the idea that migrants from poor areas, hosted by richer countries, are less integrated in their societies, and, therefore, cannot sufficiently enjoy the potential transaction cost mitigation effect of common ties.

Third, and more importantly, we still find a strong and non-linear effect of per capita GDP on remittances in income group regressions. However, while whole-sample exercises suggested a U-shaped relationship, with remittances first decreasing and then increasing with income at destination, estimates in Table 2, columns (1)–(2), show that an inverse U-shaped relationship is now in place (cf., Section 5.3 for further evidence). Therefore, contrary to what happens when one takes into account all global remittance flows, when we discriminate between origin and destination income groups, investment or exchange motives seem to be behind remittances from rich and middle countries to poor ones when home per capita GDP is low. Instead, when the income of a poor destination country grows, altruistic motives seem to dominate remittance behaviors of migrants sending money from richer countries; see Figure 3. This is in line with studies focusing on private transfers in less-developed countries (Cox et al. 1998).
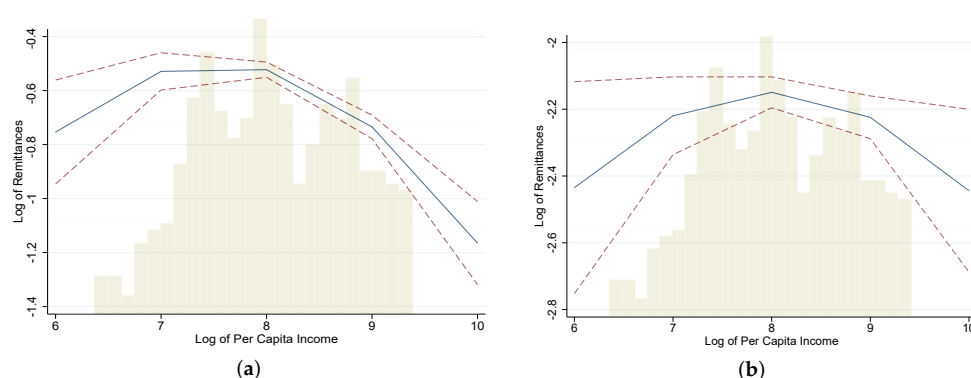
Figure 3. Non-linear marginal impact of (per capita) income (pcGDP) on remittances by income group. (**a**) Rich to poor; (**b**) middle to poor. OLS estimates from columns (1)–(2), Table 2. Solid blue line: Marginal impact. Dashed red lines: 95% confidence bands. X-axis: log of (per capita) income (pcGDP) in the observed whole-sample range of the covariate. Y-axis: log of remittance flows. The histogram in the background depicts the distribution of pcGDP at destination for countries in the poor income group across the years (bar heights are proportional to observed frequencies).

5.3. Robustness Checks

The impact of macroeconomic determinants on bilateral remittance flows estimated in the previous subsections could be biased due to a number of possibly concurrent issues. In this subsection, we attempt to address some of them in order to discuss the robustness of our foregoing findings.

To begin with, our estimates employ an originally balanced sample of 176×175 remittance flows in each year. As discussed above, this is one of the main contributions of the paper, as our aim was to have a country coverage over time as large as possible. However, the number of bilateral flows actually employed in the regressions is much smaller than the notional maximum. This is true not only in OLS estimates (where zero flows are automatically excluded) but also when PPML is employed. In both cases, the net reduction in the observations used in the estimation is due to the presence of missing values in the destination country covariates, which makes the sample actually employed for estimation unbalanced. We investigate whether this may be a source of bias in our results in a series of additional exercises in which we either shrink the sample of countries or we fit a structural gravity model where destination-country-specific covariates are replaced by country-time fixed effects (as it happens in all specifications for origin countries). In the first case, we select only the countries that do not have any missing values in all the covariates used in our baseline regressions. This allows us to focus on a smaller sample of 110 countries (see Appendix E, Table A6) out of the original 176 available. In the second case, we fit to both the reduced sample and to the 176-country sample a structural gravity specification (Baldwin and Taglioni 2006); see Equation (2). Table A7 in the Appendix E reports results for the case of baseline OLS estimates using the "stepwise" version of M_{ij}^t. We compare four different specifications, depending on whether the full 176-country sample or the reduced 110-country sample is employed, and whether the FEs are as in Equation (1) or as in the structural gravity model of Equation (2)[24]. Overall, our main results seem to be confirmed, suggesting that sample imbalance due to the presence of missing values in country covariates does not bias the main conclusions of our analysis. In particular, both the sign and magnitude of the size effect exerted by the number of migrants in the host country appears to be extremely robust in all specifications.

Our second concern is related to multicollinearity and overfitting issues. For example, population-related variables included in the baseline specification refer to the number of migrants in the host country, total population, and % of rural population at home. If these variables are strongly correlated, e.g., because the number of migrants at the origin of the remittance flows heavily depends on the population at home or its share living in rural areas, the estimated coefficients may be biased. To check for this potential source of bias, we have performed additional regressions where either the covariate "population" or the covariate "rural population share" is excluded from the list of regressors. The results are reported in the Appendix E, Table A8, columns (2)–(3), vis-à-vis the baseline correspondent estimates from Table 1 (column 1). Apart from the impact of contiguity, which becomes less relevant, all our main results robustly hold, in particular, as far as population-related regressors are concerned. Another source of multicollinearity and possible overfitting bias is related to bilateral, time-invariant variables controlling for distance, social, cultural, and political ties. These may affect remittance flows, and also migrants at the destination as well. In order to explore this potential source of bias, we have conducted two further experiments, reported in columns (9) and (10) of Table A8. In the first one (see column 9), we exclude migrants at the destination and keep bilateral variables in the regression. In the second one (see column 10), we exclude bilateral variables and keep migrants at destination in the regression. As the results in column (9) show, there is a substantial decrease in coefficient significance, although all the signs are in line with column (1). This is expected, as we are removing the only source of variation that spans the dimension ijt. When removing bilateral, time-invariant variables, we gain in coefficient significance. Notice that the coefficient of migrants at destination is substantially reduced in magnitude

and significance. Overall, however, all the correct signs are preserved and the included regressors are significant only at 5% and 1%.

Third, we investigate possible bias due to the omission of bilateral covariates controlling for technological advances, which may have contributed to reducing transaction costs. In our baseline specification, all the bilateral variables are indeed time invariant. Geographical distance alone can hardly control for technological developments that, together with an increased competition in the financial service industry, may have unevenly led to a reduction in transfer costs of remittances in particular, and of transaction costs in general (Ahmed et al. 2021; Kakhkharov et al. 2017)[25]. To address this issue, columns (4)–(8) in the Appendix E, Table A8, present estimation results where one adds to the baseline specification some proxies controlling for cost-reducing technological advances. To begin with, we have explored the impact that a larger diffusion of internet technologies at home—net of that in the host country, controlled for by origin FEs as usual—may have on remittance flows. We achieved this using two additional covariates, i.e., fixed broadband subscriptions (per 100 people) and the individuals using the internet (as a % of population); see the Appendix A, Table A1, for sources and definitions. The results in Table A8, Appendix E, columns (4)–(5), show that both variables enhance remittances, but the estimated effect of geographical distance, number of migrants, as well as those of other time-invariant bilateral variables, remain roughly in line with our baseline specification (see column (1)). Furthermore, we have exploited data on remittance costs from the World Bank "Remittance Prices Worldwide" database to build a bilateral, time-varying covariate defined as the average total cost of all remittance flows between any two countries as a percentage of the total transaction (for a similar perspective, see Ahmed et al. 2021). Unfortunately, a very small number of remittance costs are reported in that database, leading to a huge decrease in the estimation sample size. Nonetheless, the results in columns (6)–(8) of Table A8 hint at a picture that is quite consistent with our baseline regressions. More specifically, the average remittance cost appears to exert a weak negative (or a statistically insignificant) impact on remittance flows when inserted in the baseline specification with FEs as in Equation (1), together or without geographical distance[26]. Furthermore, we are still able to observe a negative (albeit weaker) effect of our contiguity covariate. In Section 5.1, this was interpreted as suggesting that sharing a border could have been an incentive to boost informal ways to remit. However, in the presence of financial development there could be a counter effect, as a more efficient financial industry could reduce remittance prices and, thus, enhance formal remittances. That counter effect was not entirely controlled for in our baseline regressions, since financial development was proxied only by bank branches at destination, whereas all bilateral covariates were time invariant. Recovering a negative effect of contiguity in this set of regressions indicates instead that sharing a border reduces formal remittances, net of the counteracting effect of financial development and technological advances. Notice that the weak impact of remittance costs is detected also when one removes time-invariant bilateral variables and replaces them with a set of paired, time-invariant FEs (Baldwin and Taglioni 2006). Indeed column (8) in Appendix E, Table A8 reports estimates when the FEs specification reads as:

$$R_{ij}^t = \kappa \exp\left\{\eta_i^t + \psi_j + \delta^t + \lambda_{ij} + \boldsymbol{\phi} M_{ij}^t + \boldsymbol{\theta} X_j^t\right\} \epsilon_{ij}^t, \tag{3}$$

where λ_{ij} is a set of time-invariant bilateral FEs. In this case, once all time-invariant transaction costs are fully controlled for, the average remittance costs weakly and negatively affect remittance flows. More importantly, however, our main results seem to be confirmed, particularly those related to the impact of migrant networks and per capita income.

Fourth, we double check whether the coefficient estimates for the migration stock variable in its stepwise version suffer from any downward bias. We replace our stepwise assumption with a migration stock variable whose non-observed values (i.e., in years 2011, 2012, 2014, 2015, and 2016) are replaced via linear interpolation of the observed values in 2010, 2013, and 2017. We obtain a 0.897 point estimate for M_{ij}^t in the OLS case (with a

standard error equal to 0.007) and a 0.846 point estimate in the PPML case (standard error equal to 0.020). This indicates that our stepwise assumption is not causing any downward bias as far as the impact of migrants at destination is concerned.

Fifth, our results may suffer from endogeneity issues, which may arise from omitted-variable biases and simultaneity, among other things. Our aim is to provide an initial robustness check against endogeneity biases, keeping, as much as possible, our analysis comparable with that performed in the existing literature studying remittances with a gravity-model approach (cf., in particular, Lueth and Ruiz-Arranz 2006; McCracken et al. 2017). Therefore, we tackle this problem with two complementary exercises; see Table A9 in the Appendix E section.[27]. First, we replace in Equation (1) possibly endogenous regressors, i.e., time-dependent country-specific covariates and migrants at destination, with their lagged (year $t - 1$) values. The results are reported in columns (1) and (2) of Table A9, depending on the assumption about migrants at destinations. As compared to our baseline estimates in Table 1, the significance seems to be weaker overall, but all signs and point estimates appear to be fairly robust. Second, we instrument each potentially endogenous regressor with its one-year lagged value and country latitude (Bugamelli and Paternò 2009; McCracken et al. 2017) and then estimate their impact on remittances in Equation (1) with panel data two-stage least squares. All first-stage F scores (testing for relevance of the instruments) are significantly different from zero and larger than 10, as required in order not to have weak instruments.[28] Furthermore, Table A9 reports outcomes from an Anderson–Rubin Wald F-test, checking for the joint significance of all endogenous regressors, and a Sargan–Hansen J test for over-identified restrictions. The results seem to confirm that the null hypothesis that all endogenous regressors are jointly equal to zero is rejected, while we cannot reject the null hypothesis that the instruments are exogenous. As to point estimates and the significance of the covariates, note that all signs and magnitudes of the coefficients are in general retained, although we can observe again a decrease in the overall significance levels.

Finally, we study whether the results about the inverse U-shaped relationship detected for the relationship between per capita GDP and remittances in the income group regressions is robust to alternative specifications. Table A10 in the Appendix E reports coefficient estimates and significance for per capita GDP and its square—for both the case of migrants at destination defined as 'Year = t' and 'stepwise'. Using Equation (1), we consider PPML estimates, the strategies employed above to control for endogeneity, our reduced 110-country sample, and possible multicollinearity induced by population-related regressors. By and large, an inverse U-shaped relationship still emerges for both "rich to poor" and "middle to poor" subsamples. However, the detected non-linearities seem to be weaker in their significance, especially when controlling for endogeneity and in the reduced 110-country sample.

6. Discussion and Conclusions

In this paper, we have explored the macroeconomic determinants of bilateral remittance flows between world countries, using the "World Bank Migration and Remittance" database, which originally covers 214 sending and receiving countries over the period 2010–2017. Exploiting the inherent origin–destination nature of remittance flows, we have fitted the data using a number of gravity-model specifications, controlling for host-, home-, and time-specific fixed effects, to a subset of 176 countries.

As discussed in Section 2, using a gravity-model approach allowed us to investigate in more detail the drivers of remittance flows, separating as much as possible the host, home, and bilateral effects. Furthermore, a gravity specification can be derived by micro-founded models (McCracken et al. 2017; Rapoport and Docquier 2006; Schiopu and Siegfried 2006), which is helpful in identifying the expected signs of coefficients in terms of migrant motives.

The results from whole-sample exercises clearly indicate that size effects (controlled for by the number of migrants at the origin- and home-country population), transaction costs (distance and contiguity), and common host–home country ties, strongly influence

global remittance flows. Furthermore, the important remittance-enhancing effect of rural population and education-related covariates hint at the existence of investment motives behind the migrant remittance behavior. We have also found that economic growth and financial development at home play an important role in impacting remittances.

Most of those macroeconomic determinants (e.g., size effects, education, economic growth, and financial development) are also important in explaining remittance flows from rich and middle countries to poor ones. However, when one conditions on the income group of host and home countries, interesting discrepancies emerge. First, the impact of transaction costs on remittances substantially changes: a higher origin–destination geographical distance between middle/rich country and poor ones boosts remittances, while sharing a border becomes less relevant. Second, common political, social, and cultural ties lose their importance in explaining remittance flows.

We have also documented the existence of a robust non-linear relationship between income at home and remittance flows. Globally, a U-shaped relationship emerges, suggesting that altruistic motives dominate when per capita GDP at destination is small, whereas self-interested or exchange motives become more relevant for higher levels of home income. On the contrary, remittance sent from middle/rich nations to poor countries are explained by self-interested motives for low-income levels and then by investment or exchange as income at home increases (i.e., an inverted U-shaped relationship between per capita GDP at destination and remittances emerges). This suggests that altruistic and self-interested motives non-trivially interact and may change across both host/home income groups and the level of income at home.

Our main results robustly hold vis-à-vis a number of alternative estimation strategies and specifications. First, as PPML-based exercises show, the most important findings are not influenced by the presence of zero-flow observations in the data. Second, coefficient estimates do not seem to be strongly affected by omitted-variable biases, since host-time fixed effects control for cross-country and longitudinal factors at the origin, destination-specific and time-invariant fixed effects control for unobserved cross-country heterogeneity at home, and year dummies for time trends. Third, we have employed two different specifications for the covariate controlling for the number of migrants, in order to mitigate the bias coming from the fact that migrant stocks are not observed in every year. Fourth, we do not detect strong departures from our main results when a number of possible additional sources of bias are considered. These include the effect of missing values in the covariates, multicollinearity between population-related regressors, endogeneity, and the presence of trends in technological advances, which may have led to remittance costs unevenly decreasing in time.

The main contribution of this paper was to employ a large panel of bilateral remittance flows among world countries in the attempt to overcome data limitations that, so far, have prevented existing studies from reaching robust and conclusive predictions on the impact of macroeconomic determinants on remittance flows (cf., Table A5 in the Appendix D). However, it must be noted that the wide cross-sectional coverage of the "World Bank Migration and Remittance" database comes at a cost. Indeed, bilateral remittance flows in the database are not empirically observed but come from an estimation procedure proposed in Ratha and Shaw (2007). As discussed in Mallela et al. (2020), remittance estimates may be inaccurate in terms of volume, especially for certain countries (Alvarez et al. 2015). Nevertheless, the database has been successfully employed in many existing works (see, e.g., Aggarwal et al. 2011; Arvin and Lew 2012; Azizi 2017). Furthermore, remittance flows are estimated using alternative weighting schemes and do not make use of gravity models, which may introduce biases in our analysis. In absence of better comprehensive data on bilateral remittance flows, this is still the best choice if one aims at a large country coverage for a sufficiently long number of years.

Exploring in more detail the possible biases that this type of remittance data may generate on gravity-model estimates is certainly one of our future avenues of research. The present work, however, may be extended in at least three additional ways.

First, the strategy employed in Lueth and Ruiz-Arranz (2006) and McCracken et al. (2017) to deal with potential endogeneity—and replicated here for the sake of comparison with the existing literature—may suffer from a number of limitations. For example, as discussed in Bellemare et al. (2017), the practice of using lagged explanatory variables, despite being commonly used, may lead to incorrect inference. Furthermore, using a time-invariant instrument in the first stage of an IV two-stage model without country fixed effects may lead to biased estimates and eventually lead to a violation of the assumption of instrument exogeneity. Therefore, more work trying to devise a robust strategy to deal with possible endogeneity in our specifications is certainly needed. Notice, however, that addressing endogeneity concerns in gravity-model exercises (no matter whether trade, migration, or other bilateral flows are concerned) is an area where, in general, the literature has not reached a consensus yet, although several alternative methodologies have been suggested (see, e.g., Bergstrand 2013; Cyrus 2002; Jochmans and Verardi 2022; Reed et al. 2016).

Second, the analysis of remittance flows disaggregated by income groups has only focused on remittances from middle and rich countries towards poor ones. Studying the behavior and determinants of other income-conditioned flows (e.g., those between rich or middle countries), as well as their geographical breakdown (e.g., north to south), may complement the present analysis.

Third, the presence and shape of non-linearities in per capita GDP (and other covariates) can be explored more deeply.

Finally, although the existing data only cover the period 2010–2017, the World Bank has recently started a new project (cf., knomad.org/data/remittances, accessed on 16 July 2023) with the goal of producing updated bilateral remittance data but, as of June 2023, only one additional matrix, for the year 2021, has been published. As new updates are made available, one could reproduce analyses in line with the present one employing more recent years.

Author Contributions: Both G.F. and T.R. contributed equally to this work. Conceptualization, G.F. and T.R.; Formal analysis, G.F. and T.R.; Investigation, G.F. and T.R.; Methodology, G.F. and T.R.; Writing—original draft, G.F. and T.R.; Writing—review & editing, G.F. and T.R.. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data employed in this work are publicly available. See Tables A1 and A2 in the Appendices A and B for a full list of data sources.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Data Sources

Table A1. Description and sources of variables used in our preferred specifications. (*) See Ratha and Shaw (2007).

Variable	Description	Data Source
Bilateral Remittances	Yearly bilateral remittance estimates, millions of US dollars (years: 2010–2017)	World Bank Migration and Remittances Data *
Number of Migrants	Bilateral estimates of migrant stocks (years: 2010, 2013, 2017)	World Bank Migration and Remittances Data *
Distance	Distance between most populated cities (km)	Cepii Gravity Database (cepii.fr)
Contiguity	Dummy variable; 1 = country pair shares a border	Cepii Gravity Database (cepii.fr)
Common Language	Dummy variable; 1 = country pair shares common official or primary language	Cepii Gravity Database (cepii.fr)
Colonial Relationship	Dummy variable; 1 = country pair ever in colonial relationship	Cepii Gravity Database (cepii.fr)
Common Religion	Dummy variable; 1 = country pair shares common religion	Cepii Gravity Database (cepii.fr)
pcGDP	Per capita GDP, PPP (constant 2011 international dollar)	World Bank Open Data (data.worldbank.org)
GDP Growth	GDP growth (annual %)	World Bank Open Data (data.worldbank.org)
Population	Population, total	World Bank Open Data (data.worldbank.org)
Rural Population Share	Rural population (% of total population)	World Bank Open Data (data.worldbank.org)
Exp on Education (% of GDP)	Government expenditure on education, total (% of GDP) pgap_550	World Bank Open Data (data.worldbank.org)
Enrollment Rate	Adjusted net enrollment rate, primary (% of primary school age children)	World Bank Open Data (data.worldbank.org)
Bank Branches Share	Commercial bank branches (per 100,000 adults)	World Bank Open Data (data.worldbank.org)
Broadband Subs	Fixed broadband subscriptions (per 100 people)	World Bank Open Data (data.worldbank.org)
Internet Usage	Individuals using the internet (% of population)	World Bank Open Data (data.worldbank.org)
Remittance Cost	Average total cost of the transaction in %	Remittance Prices Worldwide (remittanceprices.worldbank.org)
Precipitation Anomalies	Yearly total precipitation anomalies (z-score based on 1901–2018 obs)	Climatic Research Unit—CRU (www.cru.uea.ac.uk)
Temperature Anomalies	Yearly average temperatures anomalies (z-score based on 1901–2018 obs)	Climatic Research Unit—CRU (www.cru.uea.ac.uk)
Arable Land	Land cultivated for crops (% of total land area)	CIA World Factbook (www.cia.gov)
Average Elevation	Country average elevation above sea level (mt)	CIA World Factbook (www.cia.gov)
Coastline Length	Country total length of the boundary between the land area (including islands) and the sea (km)	CIA World Factbook (www.cia.gov)
Distance From The Equator	Absolute value of country latitude	CIA World Factbook (www.cia.gov)
Remoteness	Sum of distances between a country and all the others	Our own calculation based on Cepii Gravity Database (cepii.fr)

Appendix B. Additional Covariates

Table A2. Additional covariates. List of covariates used in the analysis and not included in our preferred specifications (not statistically significant in regression exercises).

Variable	Description	Data Source
Common Ethnic Language	Dummy variable; 1 = country pair shares common language (spoken by at least 9% of the population)	Cepii Gravity Database (cepii.fr)
Common Colonizer	Dummy variable; 1 = country pair shares a common colonizer post 1945	Cepii Gravity Database (cepii.fr)
Colonial Relation Post 1945	Dummy variable; 1 = country pair in colonial relationship post 1945	Cepii Gravity Database (cepii.fr)
Common Currency	Dummy variable; 1 = country pair share common currency	Cepii Gravity Database (cepii.fr)
Weighted Distance	Weighted distance (pop-wt, Km), year = 2010	Cepii Gravity Database (cepii.fr)
Domestic Credit Share	Domestic credit to private sector (% of GDP)	World Bank Open Data (data.worldbank.org)
Poverty Gap (1.90$)	Poverty gap at USD 1.90 a day (2011 PPP) (%)	World Bank Open Data (data.worldbank.org)
Poverty Gap (3.20$)	Poverty gap at USD 3.20 a day (2011 PPP) (%)	World Bank Open Data (data.worldbank.org)
Poverty Gap (5.50$)	Poverty gap at USD 5.50 a day (2011 PPP) (%)	World Bank Open Data (data.worldbank.org)
Poverty Gap Share at NPL	Poverty gap at national poverty lines (%)	World Bank Open Data (data.worldbank.org)
Educational Attainment Share	Educational attainment, at least completed primary, population 25+ years, total (%) (cumulative)	World Bank Open Data (data.worldbank.org)
Displaced Persons	Internally displaced persona, total displaced by conflict and violence (number of people)	World Bank Open Data (data.worldbank.org)
Enrollment Rate	Adjusted net enrollment rate, primary (% of primary school age children)	World Bank Open Data (data.worldbank.org)
Literacy Rate	Literacy rate, adult total (% of people ages 15 and above)	World Bank Open Data (data.worldbank.org)
Real Exchange Rate	Real effective exchange rate index (2010 = 100)	World Bank Open Data (data.worldbank.org)
Real Interest Rate	Real interest rate (%)	World Bank Open Data (data.worldbank.org)
Natural Disasters	Total number of persons affected by natural disasters	EM-DAT (www.emdat.be)
Fragility	Dummy variable; 1 = country in fragile situation (conflict, violence, and instability)	World Bank Open Data (data.worldbank.org)

Appendix C. List of Countries and Summary Statistics

Table A3. List of countries included in the baseline regression sample (176 countries).

Country	ISO3	Country	ISO3	Country	ISO3
Afghanistan	AFG	Georgia	GEO	Nicaragua	NIC
Albania	ALB	Germany	DEU	Niger	NER
Algeria	DZA	Ghana	GHA	Nigeria	NGA
Angola	AGO	Greece	GRC	Norway	NOR
Argentina	ARG	Grenada	GRD	Oman	OMN

Table A3. *Cont.*

Country	ISO3	Country	ISO3	Country	ISO3
Armenia	ARM	Guatemala	GTM	Pakistan	PAK
Australia	AUS	Guinea	GIN	Panama	PAN
Austria	AUT	Guinea-Bissau	GNB	Papua New Guinea	PNG
Azerbaijan	AZE	Guyana	GUY	Paraguay	PRY
Bahamas, The	BHS	Haiti	HTI	Peru	PER
Bahrain	BHR	Honduras	HND	Philippines	PHL
Bangladesh	BGD	Hungary	HUN	Poland	POL
Barbados	BRB	Iceland	ISL	Portugal	PRT
Belarus	BLR	India	IND	Qatar	QAT
Belgium	BEL	Indonesia	IDN	Russian Federation	RUS
Belize	BLZ	Iran, Islamic Rep.	IRN	Rwanda	RWA
Benin	BEN	Iraq	IRQ	Samoa	WSM
Bhutan	BTN	Ireland	IRL	Sao Tome and Principe	STP
Bolivia	BOL	Israel	ISR	Saudi Arabia	SAU
Bosnia and Herzegovina	BIH	Italy	ITA	Senegal	SEN
Botswana	BWA	Jamaica	JAM	Seychelles	SYC
Brazil	BRA	Japan	JPN	Sierra Leone	SLE
Brunei Darussalam	BRN	Jordan	JOR	Singapore	SGP
Bulgaria	BGR	Kazakhstan	KAZ	Slovak Republic	SVK
Burkina Faso	BFA	Kenya	KEN	Slovenia	SVN
Burundi	BDI	Kiribati	KIR	Solomon Islands	SLB
Cabo Verde	CPV	Korea, Dem. Rep.	PRK	Somalia	SOM
Cambodia	KHM	Korea, Rep.	KOR	South Africa	ZAF
Cameroon	CMR	Kuwait	KWT	Spain	ESP
Canada	CAN	Kyrgyz Republic	KGZ	Sri Lanka	LKA
Central African Republic	CAF	Lao PDR	LAO	St. Lucia	LCA
Chad	TCD	Latvia	LVA	St. Vincent and Grenadines	VCT
Chile	CHL	Lebanon	LBN	Suriname	SUR
China	CHN	Lesotho	LSO	Sweden	SWE
Colombia	COL	Liberia	LBR	Switzerland	CHE
Comoros	COM	Libya	LBY	Syrian Arab Republic	SYR
Congo, Rep.	COG	Lithuania	LTU	Tajikistan	TJK
Costa Rica	CRI	Luxembourg	LUX	Tanzania	TZA
Cote d'Ivoire	CIV	Macedonia, FYR	MKD	Thailand	THA
Croatia	HRV	Madagascar	MDG	Togo	TGO
Cuba	CUB	Malawi	MWI	Tonga	TON
Cyprus	CYP	Malaysia	MYS	Trinidad and Tobago	TTO
Czech Republic	CZE	Maldives	MDV	Tunisia	TUN
Denmark	DNK	Mali	MLI	Turkey	TUR
Djibouti	DJI	Malta	MLT	Turkmenistan	TKM
Dominica	DMA	Marshall Islands	MHL	Uganda	UGA
Dominican Republic	DOM	Mauritania	MRT	Ukraine	UKR
Ecuador	ECU	Mauritius	MUS	United Arab Emirates	ARE
Egypt, Arab Rep.	EGY	Mexico	MEX	United Kingdom	GBR
El Salvador	SLV	Micronesia	FSM	United States	USA
Equatorial Guinea	GNQ	Moldova	MDA	Uruguay	URY
Eritrea	ERI	Mongolia	MNG	Uzbekistan	UZB
Estonia	EST	Morocco	MAR	Vanuatu	VUT
Ethiopia	ETH	Mozambique	MOZ	Venezuela, RB	VEN
Fiji	FJI	Myanmar	MMR	Vietnam	VNM
Finland	FIN	Namibia	NAM	Yemen, Rep.	YEM
France	FRA	Nepal	NPL	Zambia	ZMB
Gabon	GAB	Netherlands	NLD	Zimbabwe	ZWE
Gambia, The	GMB	New Zealand	NZL		

Table A4. Descriptive statistics. Bilateral remittances and number of migrants at the origin in selected years. 176 Countries. Whole sample: all 8 years from 2010 to 2017.

	Remittances			
Year	**2010**	**2013**	**2017**	**Whole Sample**
% of Obs = 0	0.83	0.74	0.73	0.77
No. of Obs > 0	5154	8043	8035	55,731
Mean	12.50	15.61	17.18	15.27
Std Dev	208.92	244.42	282.25	249.38
Min	0.00	0.00	0.00	0.00
Max	21,693.42	22,587.29	30,019.19	30,019.19
Skewness	55.27	47.43	56.01	53.01
Kurtosis	4451.01	3237.19	4655.12	4150.52
Mean	5401.53	6559.25	6869.20	6165.94
Std Dev	88,789.56	98,966.07	96,213.84	94,942.28
Min	0.00	0.00	0.00	0.00
Max	11,600,000.00	130,00,000.00	11,600,000.00	13,000,000.00
Skewness	78.22	75.82	63.50	75.22
Kurtosis	9159.62	8989.80	6570.46	8802.70

Appendix D. Gravity Models of International Remittance Flows: Summary of the Literature

Table A5. Literature Review. Sample sizes, estimation, econometric issues, and predictions in existing papers fitting gravity models to international bilateral remittance flows. Notes: ([†]) unskilled workers only; ([‡]) number of Western Union agents; (?) sign of the prediction uncertain.

	Paper					
	Schiopu and Siegfried (2006)	**Lueth and Ruiz-Arranz (2006)**	**Docquier et al. (2012)**	**Nnyanzi (2016)**	**McCracken et al. (2017)**	**Ahmed et al. (2021)**
Sample sizes						
No. sending countries	21	16	89	African Countries	18	30
No. receiving countries	7	11	47	10	27	75
No. of years	6	25	4	21	10	7
Time period	2000–2005	1980–2004	2002–2005	1990–2011	1998–2007	2011–2017
Estimation						
Panel type	Unbalanced	Unbalanced	Unbalanced	Unbalanced	Balanced	Unbalanced
Estimation method	OLS	OLS	OLS, POISSON	OLS	OLS	OLS
Fixed effects employed	(j,t)	(i,j,t)	(i,j,t)	(t)	(t)	(i,j,t)
R^2	0.35–0.57	0.69–0.72	0.49–0.91	NR	0.72–0.92	0.46–0.70
Econometric issues						
Zero-flow treatment	No	No	Yes	No	No	No
Endogenity	No	Yes (lagged vars)	No	Yes (lagged vars)	Yes (lagged vars)	Yes (GMM)
Non-linearity in income	No	No	No	No	No	No
Rich vs. poor breakdown	No	No	Yes	No	No	No
Predictions						
Number of migrants	- [†]	+	+	+		+
Distance		-	?	-	-	?
Contiguity		-		?	+	?
Common language		+	+	?	+	?
Colonial relationship		+		?	?	?
Income (diff)	+			+	+	
Income (home)		-		-		
Income (host)		+		+		
GDP (diff)	?					
GDP (home)		+	?		+	+
GDP (host)		+	+		+	?
GDP growth (home)		?				
GDP growth (host)		-				
Real interest rate diff	?			+	?	
Inequality	?					
Remittance cost	+ [‡]					
Natural disasters (home)		?			+	
Inflation (diff)		+		+		
Credit to private sector (home)				+	+	
Credit to private sector (host)				+	-	
Unemployment				?		

Appendix E. Robustness Checks

Table A6. Countries in the Reduced Regression Sample. List of countries included in the reduced regression sample (110 countries).

Country	ISO3	Country	ISO3	Country	ISO3
Albania	ALB	Guatemala	GTM	Nicaragua	NIC
Angola	AGO	Guinea	GIN	Niger	NER
Argentina	ARG	Guinea-Bissau	GNB	Norway	NOR
Armenia	ARM	Guyana	GUY	Oman	OMN
Australia	AUS	Honduras	HND	Pakistan	PAK
Azerbaijan	AZE	Hungary	HUN	Panama	PAN
Barbados	BRB	Iceland	ISL	Paraguay	PRY
Belarus	BLR	India	IND	Peru	PER
Belgium	BEL	Indonesia	IDN	Poland	POL
Belize	BLZ	Iran, Islamic Rep.	IRN	Portugal	PRT
Benin	BEN	Ireland	IRL	Qatar	QAT
Bolivia	BOL	Israel	ISR	Russian Federation	RUS
Brazil	BRA	Italy	ITA	Rwanda	RWA
Burkina Faso	BFA	Japan	JPN	Samoa	WSM
Burundi	BDI	Kazakhstan	KAZ	Sao Tome and Principe	STP
Cabo Verde	CPV	Kenya	KEN	Senegal	SEN
Cambodia	KHM	Korea, Rep.	KOR	Sierra Leone	SLE
Cameroon	CMR	Kyrgyz Republic	KGZ	Slovenia	SVN
Chile	CHL	Lao PDR	LAO	South Africa	ZAF
Colombia	COL	Latvia	LVA	Spain	ESP
Comoros	COM	Lebanon	LBN	Sri Lanka	LKA
Costa Rica	CRI	Liberia	LBR	St. Vincent and Grenadines	VCT
Cote d'Ivoire	CIV	Luxembourg	LUX	Sweden	SWE
Croatia	HRV	Malaysia	MYS	Switzerland	CHE
Cyprus	CYP	Maldives	MDV	Tajikistan	TJK
Denmark	DNK	Mali	MLI	Tanzania	TZA
Ecuador	ECU	Malta	MLT	Togo	TGO
El Salvador	SLV	Mauritius	MUS	Tunisia	TUN
Estonia	EST	Mexico	MEX	Turkey	TUR
Ethiopia	ETH	Micronesia, Fed. Sts.	FSM	Uganda	UGA
Fiji	FJI	Moldova	MDA	Ukraine	UKR
Finland	FIN	Mongolia	MNG	United Kingdom	GBR
France	FRA	Mozambique	MOZ	United States	USA
Gambia, The	GMB	Myanmar	MMR	Uzbekistan	UZB
Georgia	GEO	Namibia	NAM	Vanuatu	VUT
Germany	DEU	Nepal	NPL	Vietnam	VNM
Ghana	GHA	New Zealand	NZL		

Table A7. Assessing the effect of missing values in the covariates. Dependent variable: bilateral remittance flows. OLS estimates of gravity-model coefficients. Columns (1) and (3): full sample (176 countries; see Appendix C, Table A3). Columns (2) and (4): reduced country sample size (110 countries; see Appendix E, Table A6). Columns (1) and (2): FEs are as in Equation (1) and destination country covariates are included. Columns (3) and (4): FEs are as in Equation (2) and destination country covariates are not included. Number of migrants: the stepwise version of the covariate is always employed (see Section 4). Significance levels: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

	(1)	(2)	(3)	(4)
Dependent Variable: Bilateral Remittance Flows	**Sample: 176 Countries**	**Sample: 110 Countries**	**Sample: 176 Countries**	**Sample: 110 Countries**
Number of Migrants (Stepwise)	0.992 *** (0.004)	0.995 *** (0.003)	0.992 *** (0.000)	0.995 *** (0.000)
Distance	−0.104 *** (0.010)	−0.091 *** (0.016)	−0.082 *** (0.001)	−0.073 *** (0.001)
Contiguity	−0.355 *** (0.035)	−0.256 *** (0.041)	−0.012 *** (0.004)	−0.023 *** (0.005)
Common Language	0.049 * (0.021)	0.037 *** (0.022)	0.039 *** (0.002)	0.049 *** (0.004)

Table A7. *Cont.*

Dependent Variable: Bilateral Remittance Flows	(1) Sample: 176 Countries	(2) Sample: 110 Countries	(3) Sample: 176 Countries	(4) Sample: 110 Countries
Colonial Relationship	0.002	0.061 ***	0.057 ***	0.0449 ***
	(0.039)	(0.032)	(0.004)	(0.006)
Common Religion	0.217 ***	0.122 **	0.111 ***	0.119 ***
	(0.025)	(0.026)	(0.003)	(0.005)
pc GDP	−1.420 ***	−1.493***	–	–
	(0.101)	(0.094)		
pc GDP Squared	0.084 ***	0.048 ***	–	–
	(0.005)	(0.005)		
GDP Growth	−0.021 ***	−0.006 ***	–	–
	(0.002)	(0.001)		
Population	0.100 ***	0.147 ***	–	–
	(0.005)	(0.007)		
Rural Population Share	0.007 ***	0.037 ***	–	–
	(0.001)	(0.001)		
Exp on Education (% of GDP)	0.075 ***	0.051 ***	–	–
	(0.004)	(0.003)		
Enrollment Rate	0.007 ***	0.0122 ***	–	–
	(0.001)	(0.002)		
Bank Branches	0.019 ***	0.009 ***	–	–
	(0.000)	(0.000)		
Origin FEs	(it)	(it)	(it)	(it)
Destination FEs	(j)	(j)	(jt)	(jt)
Destination Country Covariates	YES	YES	NO	NO
Obs	33,408	18,937	55,731	18,943
R^2	0.902	0.984	0.997	0.998
Prob > F	0.000	0.000	0.000	0.000

Table A8. Further robustness tests. OLS estimates using the full sample size (176 countries) and the stepwise version of the covariate "Number of Migrants". See Section 5.3 for a discussion. FE specification in columns (1)–(7) is as in Equation (1), in column (8) as in Equation (3). Column 1: results from the baseline specification; see Table 1. Significance levels: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Dependent Variable: Bilateral Remittance Flows	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Migrants at Destination (Stepwise)	0.992 ***	0.995 ***	0.995 ***	0.995 ***	0.995 ***	0.989 ***	0.986 ***	0.801* **	-	0.678 *
	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)	(0.029)	(0.028)	(0.011)	-	(0.297)
Distance	−0.104 ***	−0.127 ***	−0.126 ***	−0.108 ***	−0.127 ***	−0.088 **	-	-	−0.083 *	-
	(0.010)	(0.004)	(0.005)	(0.005)	(0.032)	(0.032)	-	-	(0.041)	-
Contiguity	−0.355 ***	−0.024	−0.025 *	−0.025 *	−0.022 *	−0.019 *	−0.048 *	-	−0.106 **	-
	(0.035)	(0.015)	(0.015)	(0.015)	(0.010)	(0.009)	(0.009)	-	(0.035)	-
Common Language	0.049 *	0.032 ***	0.031 ***	0.030 ***	0.031 ***	0.015 *	0.019 *	-	0.012 *	-
	(0.021)	(0.009)	(0.009)	(0.009)	(0.009)	(0.006)	(0.111)	-	(0.005)	-
Colonial Relationship	0.002	0.058 ***	0.058 ***	0.058 ***	0.058 ***	0.005	0.015	-	0.001	-
	(0.039)	(0.017)	(0.017)	(0.017)	(0.017)	(0.110)	(0.108)	-	(0.031)	-
Common Religion	0.217 ***	0.286 **	0.309 **	0.310 **	0.329 **	0.027	0.027	-	0.132 **	-
	(0.025)	(0.105)	(0.114)	(0.114)	(0.117)	(0.150)	(0.150)	-	(0.045)	-
pc GDP	−1.420 ***	−1.292 ***	−1.025 ***	−1.790 ***	−1.679 ***	−1.146 **	−1.129 **	−1.059 *	−1.001 *	−1.236 *
	(0.101)	(0.410)	(0.382)	(0.295)	(0.282)	(0.426)	(0.418)	(0.481)	(0.393)	(0.503)
pc GDP Squared	0.084 ***	0.251 ***	0.352 ***	0.438 ***	0.378 ***	0.614 **	0.605 **	0.556	0.037	0.062 **
	(0.005)	(0.029)	(0.031)	(0.032)	(0.031)	(0.306)	(0.305)	(0.360)	(0.034)	(0.022)
GDP Growth	−0.021 ***	−0.038 ***	−0.050 ***	−0.055 ***	−0.025 **	−0.016	−0.016	−0.007	−0.016 **	−0.012 *
	(0.002)	(0.012)	(0.010)	(0.011)	(0.008)	(0.010)	(0.010)	(0.011)	(0.005)	(0.006)
Rural Population Share	0.007 ***	0.052 ***	-	0.037 ***	0.033 ***	0.219 ***	0.219 ***	0.171 ***	0.003	0.005
	(0.001)	(0.006)	-	(0.006)	(0.006)	(0.053)	(0.053)	(0.058)	(0.002)	(0.004)
Exp on Education (% of GDP)	0.075 ***	0.043 ***	0.046 ***	0.039 ***	0.040 ***	0.258 ***	0.258 ***	0.219 ***	0.059 ***	0.082 **
	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)	(0.044)	(0.044)	(0.047)	(0.003)	(0.026)
Enrollment Rate	0.007 ***	0.012 ***	0.007 ***	0.013 ***	0.008 ***	0.012	0.012	0.009	0.001	0.004 *
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.009)	(0.009)	(0.010)	(0.002)	(0.002)
Bank Branches	0.019 ***	0.003 ***	0.005 ***	0.006 ***	0.007 ***	0.033 ***	0.032 ***	0.034 ***	0.015 **	0.013 **
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.011)	(0.011)	(0.011)	(0.005)	(0.004)
Population	0.100 ***	-	0.085 ***	0.053 ***	0.167 ***	0.314 ***	0.313 ***	0.073 ***	0.084 *	0.102 **
	(0.005)	-	(0.021)	(0.012)	(0.046)	(0.085)	(0.077)	(0.021)	(0.035)	(0.033)
Broadband Subs	-	-	-	0.034 ***	-	-	-	-	-	-
	-	-	-	(0.003)	-	-	-	-	-	-
Internet Usage	-	-	-	-	0.010 ***	-	-	-	-	-
	-	-	-	-	(0.001)	-	-	-	-	-
Remittance Cost	-	-	-	-	-	−0.008 *	0.010	−0.020 *	-	-
	-	-	-	-	-	(0.003)	(0.011)	(0.008)	-	-
Origin FEs	(it)	(it)	(it)	(it)	(it)	(it)	(it)	(it)	(it)	(it)
Destination FEs	(j)	(j)	(j)	(j)	(j)	(j)	(j)	(j)	(j)	(j)
Paired FEs	-	-	-	-	-	-	-	(ij)	-	-
Obs	33408	33417	33421	31015	31751	1581	1581	1581	33408	33408
R^2	0.902	0.984	0.984	0.984	0.985	0.981	0.981	0.984	0.899	0.876
Prob > F	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table A9. Endogeneity. OLS estimates. Columns (1 and 2): regressors marked by (†) are introduced with a lag in Equation (1). Columns (3 and 4): regressors marked by (†) are instrumented by their lagged observation and country latitude and estimated in Equation (1) using panel data two-stage least squares (TSLS). Null hypothesis in Anderson–Rubin Wald F-test: all coefficients of endogenous covariates are jointly equal to zero. Null hypothesis of Sargan–Hansen J test for over-identified restrictions: instruments are exogenous. Significance levels: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

	(1)	(2)	(3)	(4)
	Migrants at Destination Year = t	Migrants at Destination Stepwise	Migrants at Destination Year = t	Migrants at Destination Stepwise
Migrants at Destination †	0.887 ***	0.892 ***	0.901 ***	0.956 ***
	(0.009)	(0.011)	(0.013)	(0.014)
Distance	−0.082 ***	−0.097 ***	−0.102 **	−0.088 **
	(0.023)	(0.021)	(0.034)	(0.029)
Contiguity	−0.314 **	−0.298 ***	−0.301 **	−0.329 ***
	(0.109)	(0.043)	(0.113)	(0.062)
Common Language	0.023	0.052 *	0.028	0.043
	(0.041)	(0.026)	(0.043)	(0.026)
Colonial Relationship	0.132	0.001	0.159*	0.001
	(0.089)	(0.034)	(0.068)	(0.038)
Common Religion	0.096	0.185 ***	0.095 *	0.199 **
	(0.056)	(0.043)	(0.043)	(0.067)
pc GDP †	−1.287 **	−1.194 **	−1.329 ***	−1.097 **
	(0.411)	(0.396)	(0.386)	(0.345)
pc GDP Squared †	0.071 **	0.069 *	0.043 **	0.056 **
	(0.023)	(0.032)	(0.014)	(0.019)
GDP Growth †	−0.017**	−0.011*	−0.028 ***	−0.019 **
	(0.006)	(0.005)	(0.004)	(0.007)
Population †	0.101 *	0.076 **	0.083 ***	0.098 **
	(0.045)	(0.029)	(0.012)	(0.032)
Rural Population Share †	0.003 **	0.004 ***	0.003 **	0.002 *
	(0.001)	(0.001)	(0.001)	(0.001)
Exp on Education (% of GDP)†	0.038 *	0.074 ***	0.029 *	0.066 ***
	(0.016)	(0.009)	(0.014)	(0.012)
Enrollment Rate †	0.009 **	0.008 *	0.006	0.009 **
	(0.003)	(0.004)	(0.005)	(0.003)
Bank Branches †	0.018 **	0.012 *	0.011 **	0.014 ***
	(0.006)	(0.005)	(0.004)	(0.004)
Obs	8270	29,101	8270	29,101
R^2	0.697	0.743	0.701	0.726
Prob > F	0.000	0.000	0.000	0.000
Anderson–Rubin Wald F-Test	-	-	31.773 *** (0.000)	29.478 *** (0.000)
Sargan–Hansen J Test	-	-	1.239 (0.265)	1.373 (0.241)

Table A10. Robustness of Non-Linearity in Per Capita GDP by Income Group. Coefficient estimates for pcGDP and pcGDP squared in the two subsamples "Rich to Poor" and "Middle to Poor"; see Section 5.2. Endogeneity (1): regressors introduced with a lag. Endogeneity (2): regressors instrumented by their lagged observation and country latitude; see Table A9. Significance levels: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

		Rich to Poor		Middle to Poor	
Robustness Analysis	Migrants at Destination	Year = t	Stepwise	Year = t	Stepwise
OLS Baseline	pcGDP	1.709 *	1.644 ***	2.146 *	1.158 *
	pcGDP Squared	−0.132 **	−0.109 ***	−0.154 *	−0.073 *
PPML	pcGDP	1.478 *	1.093 **	1.623 *	1.003 *
	pcGDP Squared	−0.095 *	−0.112 *	−0.098	−0.057 *
Endogeneity (1)	pcGDP	1.612 *	1.187 *	1.489 **	1.129 *
	pcGDP Squared	−0.101 *	−0.096 *	−0.117	−0.083 *
Endogeneity (2)	pcGDP	1.543 *	1.212 *	1.512 *	1.062 *
	pcGDP Squared	−0.139	−0.089	−0.109 *	−0.090 *
Reduced Sample (110 countries)	pcGDP	1.719 **	1.348 *	1.121	1.176 *
	pcGDP Squared	−0.122 *	−0.095 *	−0.123 *	−0.129 *
Population Excluded	pcGDP	1.497 *	1.612 **	1.942 *	1.192 *
	pcGDP Squared	−0.129 **	−0.113 *	−0.133 *	−0.088 *
Rural Population Share Excluded	pcGDP	1.402 *	1.538 *	1.765 *	1.092 *
	pcGDP Squared	−0.118 *	−0.101 **	−0.121 *	−0.079 *

Notes

[1] According to the International Monetary Fund (IMF) and the World Bank (WB), remittances are defined as interpersonal transfers between migrants and their families remained in the country. They include personal transfers and compensation of employees. Throughout this paper, we will employ the terms "sending", "origin" or "host" as defining the country where migrants live and from which they send remittances; whereas "destination", "receiving" and "home" are used to qualify the country where the migrant is from.

[2] Notice, however, that remittances are typically factored in gross national product (GNP) and not in country GDP.

[3] See https://news.un.org/en/story/2019/11/1052331 (accessed on 16 July 2023) . As a result, the nominal value of remittances per migrant ballooned from about 363US$ to 2409US$.

[4] Cf., https://blogs.worldbank.org/opendata/money-sent-home-workers-now-largest-source-external-financing-low-and-middle-income (accessed on 16 July 2023).

[5] For critical surveys of this vast literature, see for example Rapoport and Docquier (2006) and Hagen-Zanker and Siegel (2007). In particular, Yang (2011) discusses the issue of migrant control over remittances, highlighting that migrants may decide to remit more the larger their control over how remittances are used in the receiving country.

[6] See, among others, Refs. Adams (2009); Freund and Spatafora (2008); Kakhkharov et al. (2017); Posso (2015); Tabit and Moussir (2016); Vargas-Silva and Huang (2006).

[7] Cf., Refs. Ahmed et al. (2021); Docquier et al. (2012); Lueth and Ruiz-Arranz (2006); McCracken et al. (2017); Nnyanzi (2016); Schiopu and Siegfried (2006). The gravity model has been the workhorse model of international trade for more than 50 years (see Baier and Standaert 2020, and references therein), but it has been successfully applied to several other bilateral-flow data, such as, e.g., equity (Portes and Rey 2005), foreign-direct investment (Harach and Rodriguez-Crespo 2014), and migration (Beine et al. 2016).

[8] We will come back to the issue of balanced vs. unbalanced estimation samples in Section 5.3.

[9] Cf., Docquier et al. (2012) for an exception.

[10] In addition, remittance cost must be sufficiently low and host-home income differential large enough.

[11] See Refs. Ahmed et al. (2021); Docquier et al. (2012); Lueth and Ruiz-Arranz (2006); McCracken et al. (2017); Nnyanzi (2016); Schiopu and Siegfried (2006).

[12] This implies that only 30 sending countries and 75 receiving countries are left in the sample, also because of the presence of many missing values in the covariate controlling for transaction costs; see also Section 5.3.

[13] For example, Schiopu and Siegfried (2006) only employ receiving-country and year FEs, whereas Lueth and Ruiz-Arranz (2006) and Docquier et al. (2012) introduce separate FEs for the sending and receiving country, as well as for years. Ref. Ahmed et al. (2021) opt instead for bilateral and year dummies only, thus neglecting cross-sectional unobservable heterogeneity at the level of sending and receiving countries. This happens also in Nnyanzi (2016), where only time FEs are considered.

[14] An alternative strategy is to use as dependent variable the ratio of remittances to the number of migrants. However, we chose not to adopt this approach as it implicitly constrains the elasticity of the stock of migrants to one and it is not usually employed in the gravity-model literature.

[15] Interpreting geographical distance as a proxy for time-invariant transaction costs is common in the gravity-model literature. After all, sheer geography should have largely decreased its impact on international bilateral flows in the era of globalization, not only in the case of trade or migration, but also when immaterial goods are concerned (Coe et al. 2007). However, geographical distance still appears to be a large and growing obstacle to bilateral flows even when it proxies immaterial transport costs (Brei and von Peter 2018).

[16] See Section 4 and Tables A1 and A2 in the Appendices A and B for more details.

[17] We have also experimented with specifications where country GDP instead of population is used to proxy country size, without any substantial differences in our results.

[18] A positive effect of education on remittances is also found by Bollard et al. (2011) using using microdata in 11 major host countries

[19] Unfortunately, data before 2010 and after 2017 are not available, although the World Bank has recently launched a new project aiming at publishing updated bilateral remittance data. At the time of writing, only one additional matrix for year 2021 is available, cf., knomad.org/data/remittances (accessed on 16 July 2023).

[20] Indeed, the HT test without trend returns a z-score of -1.3641 (p-value= 0.0863), whereas the HT test with trend returns a z-score of -6.3988 (p-value = 0.0000). Note that z-score of the HT statistic is asymptotically $N(0,1)$ distributed.

[21] Also in this case, the explaining power of many additional, potentially interesting, factors has been explored. Due to their non-significance in most of the regression exercises performed, they have been excluded form our preferred specification. These additional regressors are: domestic credit share, poverty gaps, educational attainment, enrollment and literacy rates, the number of displaced persons, real exchange and interest rates, intensity of natural disasters, and country-fragility indicators (see Appendix B, Table A2 for details)

[22] See https://datahelpdesk.worldbank.org/knowledgebase/articles/906519 (accessed on 16 July 2023)).

23 Overall, the main insights from Table 2 robustly hold using PPML estimators.

24 Some missing values are obviously present in dyadic time-invariant covariates, but we decided not to shrink the country sample size further to keep a sufficiently large country coverage

25 It must be noted that our only time-varying bilateral variable (i.e., "Number of Migrants") might partly control for differences in transaction costs, as the corridors with a larger number of migrants and higher competition tend to exhibit consistently lower remittance costs.

26 This is partly in contrast with findings in Ahmed et al. (2021), who find a strong negative impact of remittance costs in absence of geographical distance. Such discrepancy may be due to a number of reasons. First, we employ a richer FEs specification to control for origin and destination unobserved heterogeneity and a larger set of covariates. Second, the definition of the average cost of remittances somewhat differs, as Ahmed et al. (2021) build a covariate computing the cost sending USD200 as a percentage of the amount remitted.

27 We refer the reader to Section 6 for a discussion of potential pitfalls of the strategies employed and possible extensions

28 This is the rule of thumb suggested by Stock and Staiger (1997). F-tests results are not reported in Table A9 but they are available from the Authors upon request.

References

Adams, Richard. 2009. The determinants of international remittances in developing countries. *World Development* 37: 93–103. [CrossRef]

Aggarwal, Reena, Asli Demirguc-Kunt, and Maria Soledad Martínez Peria. 2011. Do remittances promote financial development? *Journal of Development Economics* 96: 255–64. [CrossRef]

Ahmed, Junaid, Mazhar Mughal, and Inmaculada Martínez-Zarzoso. 2021. Sending money home: Transaction cost and remittances to developing countries. *The World Economy* 44: 2433–59. [CrossRef]

Alvarez, Sandra Paola, Pascal Briod, Olivier Ferrari, and Ulrike Rieder. 2015. Remittances: How reliable are the data? *Migration Policy Practice* 5: 1–5.

Anderson, James E., and Eric van Wincoop. 2003. Gravity with gravitas: A solution to the border puzzle. *American Economic Review* 93: 170–92. [CrossRef]

Arvin, Mak, and Byron Lew. 2012. Do happiness and foreign aid affect bilateral migrant remittances? *Journal of Economic Studies* 39: 212–30. [CrossRef]

Azizi, Seyed Soroosh. 2017. Altruism: Primary motivation of remittances. *Applied Economics Letters* 24: 1218–21. [CrossRef]

Baier, Scott, and Samuel Standaert. 2020. Gravity models and empirical trade. In *Oxford Research Encyclopedia of Economics and Finance*. Oxford: Oxford University Press.

Baldwin, Richard, and Daria Taglioni. 2006. *Gravity for Dummies and Dummies for Gravity Equations*. Working Paper 12516, September. Cambridge: National Bureau of Economic Research.

Beine, Michel, Simone Bertoli, and Jesus Fernandez-Huertas Moraga. 2016. A practitioners guide to gravity models of international migration. *The World Economy* 39: 496–512. [CrossRef]

Bellemare, Marc F., Takaaki Masaki, and Thomas B. Pepinsky. 2017. Lagged explanatory variables and the estimation of causal effect. *The Journal of Politics* 79: 949–63. [CrossRef]

Bergstrand, Jeffrey H. 2013. Measuring the Effects of Endogenous Policies on Economic Integration. *CESifo Economic Studies* 59: 199–222. [CrossRef]

Bollard, Albert, David McKenzie, Melanie Morten, and Hillel Rapoport. 2011. Remittances and the brain drain revisited. *World Bank Economic Review* 25: 132–56. [CrossRef]

Brei, Michael and Goetz von Peter. 2018. The distance effect in banking and trade. *Journal of International Money and Finance* 81: 116–37. [CrossRef]

Bugamelli, Matteo and Francesco Paternò. 2009. Do workers' remittances reduce the probability of current account reversals? *World Development* 37: 1821–38. [CrossRef]

Carling, Jørgen. 2008. The determinants of migrant remittances. *Oxford Review of Economic Policy* 24: 581–98. [CrossRef]

Coe, David, Arvind Subramanian, and Natalia Tamirisa. 2007. The missing globalization puzzle: Evidence of the declining importance of distance. *IMF Staff Papers* 54: 34–58. [CrossRef]

Cooray, Arusha, and Debdulal Mallick. 2013. International business cycles and remittance flows. *The B.E. Journal of Macroeconomics* 13: 515–47. [CrossRef]

Cox, Donald, Zekeriya Eser, and Emmanuel Jimenez. 1998. Motives for private transfers over the life cycle: An analytical framework and evidence for peru. *Journal of Development Economics* 55: 57–80. [CrossRef]

Cyrus, Teresa L. 2002. Income in the gravity model of bilateral trade: Does endogeneity matter? *The International Trade Journal* 16: 161–80. [CrossRef]

de Sousa, Jose, and Laetitia Duval. 2010. Geographic distance and remittances in romania: Out of sight, out of mind? *International Economics* 121: 81–97. [CrossRef]

Docquier, Frederic, Hillel Rapoport, and Sara Salomone. 2012. Remittances, migrants' education and immigration policy: Theory and evidence from bilateral data. *Regional Science and Urban Economics* 42: 817–28. [CrossRef]

Freund, Caroline, and Nikola Spatafora. 2008. Remittances, transaction costs, and informality. *Journal of Development Economics* 86: 356–66. [CrossRef]

Giuliano, Paola, and Marta Ruiz-Arranz. 2009. Remittances, financial development, and growth. *Journal of Development Economics* 90: 144–52. [CrossRef]

Hagen-Zanker, Jessica, and Melissa Siegel. 2007. *The Determinants of Remittances: A Review of the Literature*. Technical Report, MGSoG Working Paper. Available online: https://ssrn.com/abstract=1095719 (accessed on 16 July 2023).

Harach, Monika, and Ernesto Rodriguez-Crespo. 2014. *Foreign Direct Investment and Trade: A Bi-Directional Gravity Approach*. Kiel Advanced Studies Working Papers 467. Kiel: Kiel Institute for the World Economy (IfW).

Harris, Richard, and Elias Tzavalis. 1999. nference for unit roots in dynamic panels where the time dimension is fixed. *Journal of Econometrics* 91: 5201–226. [CrossRef]

Jochmans, Koen, and Vincenzo Verardi. 2022. Instrumental-variable estimation of exponential-regression models with two-way fixed effects with an application to gravity equations. *Journal of Applied Econometrics* 37: 1121–37. [CrossRef]

Kakhkharov, Jakhongir, Alexandr Akimov, and Nicholas Rohde. 2017. Transaction costs and recorded remittances in the post-soviet economies: Evidence from a new dataset on bilateral flows. *Economic Modelling* 60: 98–107. [CrossRef]

Lucas, Robert E. B., and Oded Stark. 1985. Motivations to remit: Evidence from botswana. *Journal of Political Economy* 93: 901–18.

Lueth, Erik, and Marta Ruiz-Arranz. 2006. *A Gravity Model of Workers' Remittances*. IMF Working Papers 2290. Washington, DC: International Monetary Fund.

Mallela, Keerti, Sunny Kumar Singh, and Archana Srivastava. 2020. Estimating bilateral remittances in a macroeconomic framework: Evidence from top remittance-receiving countries. *Studies in Microeconomics* 8: 95–118. [CrossRef]

McCracken, Scott, Carlyn Ramlogan-Dobson, and Marie M Stack. 2017. A gravity model of remittance determinants: evidence from latin america and the caribbean. *Regional Studies* 51: 737–49. [CrossRef]

Nnyanzi, John Bosco. 2016. What drives international remittances to africa: Altruism, self-interest or the institutional environment? *African Journal of Economic and Management Studies* 7: 397–418. [CrossRef]

Portes, Richard, and Helene Rey. 2005. The determinants of cross-border equity flows. *Journal of International Economics* 65: 269–96. [CrossRef]

Posso, Alberto. 2015. Remittances and financial institutions: is there a causal linkage? *The BE Journal of Macroeconomics* 15: 769–89. [CrossRef]

Rapoport, Hillel, and Frederic Docquier. 2006. The economics of migrants' remittances. In *Handbook of the Economics of Giving, Altruism and Reciprocity* . Edited by Serge-Christophe_Kolm and Jean Mercier_Ythier. Amsterdam: Elsevier, vol. 1, Chapter 17, pp. 1135–98.

Ratha, Dilip, and William Shaw. 2007. *South-South Migration and Remittances*. World Bank Working Paper 102. Washington, DC: The World Bank.

Reed, Robert, Christina Lira, Lee Byung-Ki, and Junsoo Lee. 2016. Free trade agreements and foreign direct investment: The role of endogeneity and dynamics. *Southern Economic Journal* 83: 176–201. [CrossRef]

Santos Silva, João, and Silvana Tenreyro. 2006. The log of gravity. *The Review of Economics and Statistics* 88: 641–58. [CrossRef]

Schiopu, Ioana C, and Nikolaus Siegfried. 2006. *Determinants of Workers' Remittances: Evidence from the European Neighbouring Region*. ECB Working Paper 688. Frankfurt: European Central Bank.

Stark, Oded, and You Qiang Wang. 2002. Migration dynamics. *Economics Letters* 76: 159–64. [CrossRef]

Stock, James, and Douglas Staiger. 1997. Instrumental variables regression with weak instruments. *Econometrica* 65: 557–86.

Tabit, Safaa and Charaf-Eddine Moussir. 2016. Macroeconomic determinants of migrants remittances: Evidence from a panel of developing countries. *International Journal of Business and Social Research* 6: 1–11. [CrossRef]

Vargas-Silva, Carlos, and Peng Huang. 2006. Macroeconomic determinants of workers' remittances: Host versus home country's economic conditions. *The Journal of International Trade and Economic Development* 15: 81–99. [CrossRef]

Yang, Dean. 2011. Migrant remittances. *Journal of Economic Perspectives* 25: 129–52. [CrossRef]