

Governance Quality and Net Migration Flows

Andrea Ariu*

Frédéric Docquier[†]

Mara P. Squicciarini[‡]

Abstract

This paper shows that governance quality promotes positive *net* inflows of high-skilled migrants. Home and foreign institutions influence both inflows and outflows, thus determining the net flows of college graduate migrants. Therefore, institutions can affect human capital through migration flows. Our empirical strategy is based on a random utility model from which we derive the net balance of migrants and an exclusion restriction to control for the selection of migrants. We test the predictions of the model using comprehensive matrices of migration by education level and a synthetic indicator of governance quality. We account for endogeneity concerns by means of an instrumental strategy and we disentangle the effect of the quality of domestic and foreign institutions on both inflows and outflows.

*University of Geneva, 40 Bv du Pont d'Arve, 1205 Geneva, Switzerland. McDonough School of Business, Georgetown University, USA. CRENOS, Italy. Email: andrea.ariu@unige.ch. Tel. +41223798264.

[†]FNRS and IRES, Université catholique de Louvain, Place Montesquieu 3, 1348 Louvain la Neuve, Belgium. Email: frederic.docquier@uclouvain.be Tel. +3210474149.

[‡]FWO and LICOS, KULeuven, Waaistraat 6, 3000 Leuven, Belgium. Email: mara.squicciarini@kuleuven.be Tel. +3216326591.

1 Introduction

In an era when human capital is crucial for economic growth, factors attracting foreign workers are as important as those refraining natives from emigrating. Both immigration and emigration shape net migration flows and affect the human capital accumulation process.

In this paper, we investigate the effect of governance quality on *net* flows of human capital defined as differences between inflows and outflows of migrants by education level.¹ We find that the quality of institutions has a positive effect on the net inflow of college-educated migrants. In particular, college graduates are more willing to migrate to countries with good institutional quality, and they tend to emigrate more from countries with low governance quality despite potentially greater migration costs. The results for the less educated have a slightly different pattern. More precisely, the low-skilled are also more likely to leave countries with low institutional quality, but we find no effect of the quality of foreign institutions on their migration choices. Therefore, the difference in institutional quality between the home country and the destination country seems to be less important to explain the net migration flows of low-skilled migrants compared to high-skilled migrants.

The analysis of the paper proceeds in three steps. First, we develop a theoretical framework and use a random utility model of migration that delivers migration balances as a function of bilateral differences in country characteristics. The model predicts that net migration flows are a function of asymmetries in the quality of institutions, wages, population size, and diasporas. The main advantage of focusing on net migration flows rather than on unilateral emigration or immigration flows separately is that all symmetric factors (observed or unobserved) affecting both immigration and emigration cancel each other out. Therefore, the model itself provides a rationale for an exclusion restriction to control for the selection of migrants. In addition, focusing on net flows reduces the scope of omitted variable problems. Any unobserved bilateral factor that influences immigration and emigration symmetrically, such as cultural proximity, does not affect net flows.

In the second step, we provide descriptive evidence of the model by correlating net flows derived from comprehensive matrices of migration (Artuc, Docquier, Ozden, and Parsons, 2015) and a synthetic indicator of governance quality derived from the six governance measures provided by Kaufmann, Kraay, and Mastruzzi (2009). We deal with selection on inflows and outflows by following the strategy of Helpman, Melitz, and Rubinstein (2008) and, as predicted by the model, we use the symmetric bilateral components of migration costs as exclusion restrictions. The results show a positive correlation between governance quality and net migration

¹In the analysis, we take into account the possible imperfect substitutability between emigrants and immigrants.

flows and provide descriptive evidence for the predictions of the model. However, these regressions can suffer both from an omitted variable and a reverse causality problem. The first can result from unobserved asymmetric factors that influence both net migration flows and the quality of institutions. The second can be due to: i) immigrants directly influencing the institutions of the host country by voicing their opinion and voting (Hirschman, 1970); ii) emigration rates increasing the incentives for the elite to improve the quality of institutions (Docquier and Rapoport, 2003); and iii) emigrants voicing their opinion from abroad (Li and Hale, 2005; Spilimbergo, 2009; Docquier, Lodigiani, Rapoport, and Schiff, 2016).² Thus, a sound instrumental strategy is needed to capture the causal effect. We instrument the distance in the quality of governance between two countries with the distance in the Scrabble index of their name. Language traits influence the set of norms and values which constitute institutions in a country (Tabellini, 2008). We show that the quality of institutions is negatively correlated with the complexity of a country's name. Therefore, countries with more complex languages - and thus more complex names - tend to have worse institutions.³ At the same time, we show that migrants do not take into account the complexity of a country's name when migrating. Therefore, the instrument is correlated with the endogenous variable and is orthogonal to migration flows. Both the first and second stages of our 2SLS strategy perform as expected and allow us to identify the positive and significant impact of governance quality on migration balances for college graduates.

Finally, in the third step, we disentangle the positive effect of the quality of institutions on net migration flows by looking separately at inflows and outflows and by separating the effect of home and foreign institutions. We find that college graduates take into account both home and foreign institutional quality when choosing where to migrate, while the low-skilled only consider home institutional quality. This can be the result of the low-skilled having more trouble acquiring and/or processing information on foreign countries. The insignificant elasticity of the low-skilled with respect to foreign institutions helps explain why the difference in institutional quality between home and foreign countries is less important for low-skilled than for high-

²Li and Hale (2005) were the first to provide a cross-country investigation of the impact of skilled labor migration on a sending country's institutional development. Spilimbergo (2009) found that foreign-trained students promote democracy in their home countries only if the foreign education was acquired in a democratic country. More recently, Docquier et al. (2016) found a robust and positive effect of emigration on the quality of institutions in a panel setting.

³Tabellini (2008) argues that language can influence the culture traits at the basis of institutions and their quality, such as the general morality of the people living in a country, defined as "*the universal applicability of rules of just conduct*". At the same time, Chen (2013) shows that linguistic traits affect both economic and non-economic attitudes. On the one hand, a more complex language might have led to more difficulties in coordination and comprehension, thus limiting the development of institutions. On the other hand, it might have contributed to developing more complex and thus less efficient institutions (Room, 2015).

skilled net migration flows.

This paper contributes to an increasing segment of the literature on the determinants of international migration. Previous work studied the determinants of bilateral migration stocks and flows (e.g. Belot and Hatton, 2012; Mayda, 2010; Grogger and Hanson, 2011; Beine, Docquier, and Ozden, 2011; Bertoli and Fernández-Huertas Moraga, 2015, 2013), or aggregate immigration and emigration flows (e.g. Pedersen, Pytlikova, and Smith, 2008; Docquier, Lohest, and Marfouk, 2007). Our contribution to this literature is twofold. First, we focus on the determinants of the size and skill structure of *net* migration flows (i.e. differences between inflows and outflows by education level). Second, while previous studies analyzed the role of income (Belot and Hatton, 2012; Grogger and Hanson, 2011), migrants' networks (Beine et al., 2011), or migration policies (Bertoli and Fernández-Huertas Moraga, 2013), we focus on the role of governance quality. However, in comparative growth studies, the quality of institutions has been considered by some influential economists as a major explanation of cross-country inequality (e.g. Hall and Jones, 1999; Acemoglu, Johnson, and Robinson, 2005; Acemoglu, Johnson, Robinson, and Yared, 2005; Shleifer, de Silanes, and La Porta, 2008). Hence, it is worth investigating whether the effect of institutions on growth is partly channelled through the mobility of highly educated workers and less educated ones.

The remainder of the paper is organized as follows. Section 2 explains the micro-foundations of our empirical strategy. The data are described in Section 3. Section 4 presents the empirical results. Finally, Section 5 concludes.

2 Empirical Strategy

Our empirical strategy is based on a random utility model of migration, which provides a rationale for an exclusion restriction to control for the selection of migrants.

Random utility model. Individuals born in an origin country i ($i = 1, \dots, I$) decide whether to stay in their home country or emigrate to another country j ($j = 1, \dots, J$). For simplicity, we abstract from skill heterogeneity, but our micro-foundations could be made specific to a particular skill, age, or gender group. The indirect utility of an individual is linear in income (as in Grogger and Hanson (2011)), in the quality of institutions, and includes possible migration costs.

In a given group, the utility of an individual born in country i and staying in country i is given by

$$u_{ii} = \alpha w_i + \beta G_i + \varepsilon_{ii} \equiv \bar{u}_{ii} + \varepsilon_{ii}$$

where w_i denotes the expected labor income in location i , G_i denotes the quality of governance

and institutions, ε_{ii} is a spatially uncorrelated individual-specific iid random term;⁴ we assume ε_{ii} follows a type-I extreme-value distribution. Coefficient α measures the marginal utility of income; and β denotes the preference for staying in a country with good institutions. The utility obtained when the same person migrates to location j is given by

$$u_{ij} = \alpha w_j + \beta G_j - C_{ij} + \varepsilon_{ij} \equiv \bar{u}_{ij} + \varepsilon_{ij}$$

where w_j , G_j and ε_{ij} denote the same variables as above, and C_{ij} captures moving and assimilation costs that are borne by the migrant. Here, coefficient β captures the preference for the quality of institutions at destination. When the random term follows an iid extreme-value distribution, we can apply the results of [McFadden \(1984\)](#) and write the log ratio of emigrants in country j to residents of i as:

$$\ln \left[\frac{M_{ij}}{M_{ii}} \right] = \alpha [w_j - w_i] + \beta [G_j - G_i] - C_{ij} \quad (1)$$

Migration costs are not observable. In line with the rest of the literature, we assume they increase with the distance d_{ij} between the two countries (i.e. geographical, cultural, and linguistic distances), decrease with the size of the established migration network or diaspora N_{ij} , decrease with the size of the native population in the host country M_{jj} (a country's capacity to host migrants increases with the size of the native population), and decrease with the quality of institutions at origin G_i . These effects are likely to vary across groups. Using the logarithmic, we write

$$C_{ij} = \delta \ln d_{ij} - \epsilon \ln N_{ij} - \rho \ln M_{jj} - \lambda G_i \quad (2)$$

where coefficient λ captures the fact that bad institutions and low government effectiveness at origin can be responsible for greater emigration costs. It is well known that emigrating from Socialist countries is very costly; think about the thousands of Cubans taking to the sea to reach the United States on makeshift crafts. More generally, [McKenzie \(2007\)](#) collected data on passport costs in 127 countries and showed that they represent more than 10 percent of the annual income per capita in 13 countries (e.g. 125 percent in the Democratic Republic of Congo, 60 percent in Chad, 51 percent in Burundi, etc.). High passport costs are found to be associated with poor governance, especially in terms of the quality of the bureaucracy, and with lower levels of migration. Moreover, skill differences in retention policies can be implemented by governments aiming at preventing the brain drain.

⁴See [Bertoli and Fernández-Huertas Moraga \(2013\)](#) for a relaxation of this hypothesis.

Migration outflows. Plugging (2) into (1) gives the expression for the bilateral migration outflow from i to j :

$$\ln M_{ij} = \alpha (w_j - w_i) + \beta G_j - (\beta - \lambda) G_i - \delta \ln d_{ij} + \epsilon \ln N_{ij} + \rho \ln M_{jj} + \ln M_{ii}$$

Clearly, the bilateral migration outflow is an asymmetric function of governance quality at destination and origin. The coefficient for governance at destination only captures the preference channel, while the coefficient for governance at origin captures the joint effects of preferences and migration costs. When this expression is negative, it means that less than one migrant is willing to move from i to j . This means a zero migration flow is observed, a frequent situation in the data since there is none or very little migration between some distant country pairs.⁵

Migration balances. For each pair of countries, we define a migration balance as the log ratio of immigrants to emigrants. Combining expressions for $\ln M_{ij}$ and $\ln M_{ji}$ from (3), the structural equation for the migration balance is the following:

$$\ln \frac{M_{ji}}{M_{ij}} = 2\alpha (w_i - w_j) + (2\beta - \lambda) (G_i - G_j) + \epsilon \ln \frac{N_{ji}}{N_{ij}} + (1 - \rho) \ln \frac{M_{jj}}{M_{ii}}$$

It shows that the migration balance improves with the wage gap between countries, the difference in institutional quality, the log-ratio of established diasporas, and the log-ratio of stayers' populations. Interestingly, the bilateral migration balance only depends on the difference in governance quality between the destination and origin countries, and the coefficient for this variable jointly accounts for the preference for good governance (2β), and the effect of governance quality on emigration costs (λ).

We cannot estimate (3) using OLS given the high occurrence of zeroes in both inflows and outflows. In addition, we cannot use the PPML estimator because some net flows have negative values. We proceed similarly to [Helpman et al. \(2008\)](#) who analyzed net trade flows across country pairs. We first run a Probit regression separately for inflows and outflows. In both cases, the dependent variable takes value one if we observe a positive flow and zero otherwise. The independent variables are the same as in equation (3), plus the exclusion restrictions. Our theoretical micro-foundations show that symmetric variables (geographic, linguistic, and cultural distances, d_{ij}) balance out in (3) and do not affect the size of migration balances. We use these bilateral symmetric characteristics to predict the probability that both $\ln M_{ij}$ and $\ln M_{ji}$ are positive. In other words, the random utility model itself provides the rationale for the use

⁵In practice, some reported zeroes might not reflect the actual absence of migrants. Due to confidentiality and disclosure rules, some statistics offices report a zero when the diaspora size is below a threshold value.

of exclusion restrictions in the empirical strategy. From these Probit regressions, we compute the Mills ratios for both inflows and outflows ($\ln R_{ji}$ and $\ln R_{ij}$), and we take their difference ($\ln R_{ji} - \ln R_{ij}$). In [Helpman et al. \(2008\)](#), the two Mills ratios control for the probability of observing zero import or export flows. Their interpretation is that none of the potential exporting firms can pay the fixed cost of exporting. Similarly, in our paper, the two Mills ratios control for the probability that none of the potential emigrants or immigrants find the benefits of migrating higher than the cost. In the second stage, we run an OLS regression as in (3), augmented with the difference in the Mills ratios.

3 Data and Stylized Facts

The core of our analysis is based on the estimation of equation (3) for which we need six inputs: migration flows (both inflows and outflows), wages (for both high-skilled and low-skilled migrants), size of migration networks, size of resident populations, quality of governance, and geographical variables as exclusion restrictions. In this section, we explain our data sources and present some stylized facts and descriptive statistics.

3.1 Data Sources

Migration data. The migration data were obtained from [Artuc et al. \(2015\)](#) who produced 195x195 comprehensive matrices of bilateral migration stocks. These matrices are computed for two skill groups (college graduates and less educated individuals) and for two years (1990 and 2000). Migration is defined on the basis of the country of birth. The dataset only includes people aged 25 and over as a proxy for the working-age population. This excludes a large number of students who emigrate temporarily to complete their education, or children who migrate with their families and are not yet active in the labor market. The methodology used in [Artuc et al. \(2015\)](#) consists of three steps. The starting point is the database described in [Docquier, Lowell, and Marfouk \(2009\)](#) documenting bilateral migration stocks to OECD host countries. It is based on a collection of census and register immigration data by country of birth and educational level for 30 OECD countries. The second step consists of a collection of similar immigration data from 46 non-OECD destinations in 2000, and 30 countries in 1990. Finally, the data collected in steps 1 and 2 are used to predict the size and structure of migration to the remaining 119 non-OECD host countries in 2000 and 135 countries in 1990.

The relevant migration variables to be used in our analysis are the log net migration flows, $\ln \frac{M_{ij}}{M_{ji}}$, and the log of the diasporas, $\ln \frac{N_{ij}}{N_{ji}}$. The net flow of immigrants to a country can be recovered by taking the difference of the logged inflows (M_{ij}) and outflows (M_{ji}). Both inflows and outflows are computed as the change in bilateral stocks of migrants between 1990 and

2000, adjusted for the average mortality rate of individuals aged 65 and over in the destination countries. The construction of destination-specific mortality rates is based on life expectancy data taken from the World Development Indicators.⁶ Since we are in a net setting, we observe the same country pair twice (e.g. both Italy-Belgium and Belgium-Italy). In this way, the sample contains the same net migration flow twice (but with opposite signs). To avoid inflating our estimation sample and having errors correlated within the same country pair, we only consider each country pair once. The two diasporas, N_{ij} and N_{ji} , are simply measured as the stocks of foreign-born people in a destination country i from a certain origin country j in 1990. Since the net balance is computed separately for each of the two educational levels, our diaspora measures are also skill specific. We believe that when migrating to a foreign country, skill-specific diasporas are more effective than the total diaspora in attracting migrants from the same education group because of common interests and affinity.

The data on the resident national populations are also from [Artuc et al. \(2015\)](#). Variables M_{ii} and M_{jj} are computed as the difference between the resident and immigrant populations in 2000 and in 1990. Migration balances, diasporas, and stayers' populations are computed both for college graduates and for less educated individuals.

Workers' relative productivity by origin. It is widely documented that many immigrants with higher education tend to find jobs in occupations typically staffed by less educated natives (see [Mattoo, Neagu, and Özden \(2008\)](#)). In particular, highly educated immigrants trained in developing countries may be less productive in high-skill jobs than natives with similar educational degrees. Evidence of such heterogeneity in the quality of education is provided by [Coulombe and Tremblay \(2009\)](#), who compare the skill intensity and schooling levels of Canadian immigrants and natives who were both submitted to standardized tests in literacy, math, and problem-solving. These tests provide measures of proficiency that are comparable across countries and over time. On this basis, [Coulombe and Tremblay \(2009\)](#) estimate a "skill-schooling gap" expressed in years of schooling. A skill-schooling gap of n years for a given country means that Canadian nationals with y years of schooling are as productive as immigrants with $y + n$ years of schooling. The greater the skill-schooling gap, the lower the quality of education in the country of origin. Simple bivariate OLS regressions show that the skill-schooling gap is a decreasing function of the per worker income in the origin country. Their -0.10 point estimate of the slope coefficient indicates that the skill-schooling gap is one year smaller when

⁶Several papers, such as [Beine et al. \(2011\)](#) and [Bertoli and Fernández-Huertas Moraga \(2015\)](#), used a similar technique to obtain emigration and/or immigration flows. We are aware that computing migration flows using the adjusted difference in stocks is not optimal, since mortality differences between natives and migrants and return migration can introduce noise. However, this is the only available alternative in the absence of actual data on migration flows by education level.

the per worker income increases by USD 10,000 in the origin country. Using this estimate and cross-country data on per worker income, we construct an indicator of the skill-schooling gap for each origin country. Then, assuming that one year of schooling generates a productivity gain of 8 percent, we estimate the relative productivity of educated immigrants and natives in each country, with a benchmark value of one for workers trained in Canada (as well as workers trained in richer origin countries, i.e. the upper bound of this index is one). For example, college graduate immigrants from Angola and Portugal have productivity levels equal to 73 and 85 percent of Canadian college graduates, respectively. Given this heterogeneity in educational quality among countries, we use adjusted migration balances in our benchmark specification. For each origin country, we weight high-skilled migration flows by the relative productivity of college-educated emigrants to Canada, thus accounting for the imperfect comparability between inflows and outflows. In a robustness analysis, we consider unadjusted migration balances by simply differentiating exits from entries.

Wage rates by education level. In many empirical studies, wages are approximated by the level of GDP per worker (e.g. [Mayda \(2010\)](#)). We need wage data here for the two education levels (as in [Grogger and Hanson \(2011\)](#), and [Belot and Hatton \(2012\)](#)). To identify skill-specific wage rates, we need data on average wages, on the skill composition of the labor force, and on wage disparities across groups. Assuming a constant share of labor income of 70 percent in all countries, the average wage of workers is assumed to be equal to 70 percent of the GDP per worker. By definition, this average wage is a weighted sum of high-skilled and low-skilled wages. The data on GDP per worker in USD in 1990 and 2000 are obtained from the World Development Indicators. The data on the education structure of the resident labor force are obtained from [Artuc et al. \(2015\)](#). As for the wage ratios between college graduates and the less educated, we combine data on returns to schooling and average years of schooling by educational attainment. Mincerian returns to schooling, MR , are available for 54 countries in [Hendricks \(2004\)](#). For the same countries, we use [Barro and Lee \(2010\)](#) data and compute the difference in years of schooling, DY , between college graduates and the less educated. The wage ratio is then computed as $(1 + MR)^{DY}$. The resulting wages are expressed in thousands of USD.

Quality of governance. We use data from the "Governance Matters" project, started with the seminal work of [Kaufmann, Kraay, and Zoido-Lobaton \(1999\)](#).⁷ The database reports six broad dimensions of governance – ranging from approximately -2.5 to 2.5 – for over 200 countries over the period 1996-2011: (i) Voice and Accountability captures perceptions of the extent

⁷The most recent methodology is described in [Kaufmann et al. \(2009\)](#).

to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and free media; (ii) Political Stability and Absence of Violence measures perceptions of the likelihood that the government will be destabilized or overthrown by unconstitutional or violent means, including politically-motivated violence and terrorism; (iii) Government Effectiveness captures perceptions of the quality of public services, the quality of the civil service, and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies; (iv) Regulatory Quality captures perceptions of the ability of the government to formulate and implement sound policies and regulations that allow and promote private sector development; (v) Rule of Law captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence; and (vi) Control of Corruption captures perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as the "capture" of the state by elites and private interests.

These six dimensions of governance exhibit pairwise correlation rates from 0.60 to 0.95, with a mean of 0.85. For this reason, it is very difficult to identify the dimensions that induce the largest push/pull effects. We thus perform a principal component analysis (PCA) to reduce the dimension of the governance indicators. Using an orthogonal transformation, the PCA creates an artificial variable which accounts for as much of the variability observed in the data. In our analysis, we use the standardized first principal component of the six governance variables which will be referred to as PCA. We use PCA as a synthetic indicator of governance; it accounts for 88% of the total variability in governance quality.

Distance variables. Distance variables (geographical, cultural, and linguistic variables) are usually considered important predictors of migration flows. Because of their symmetric nature, they represent our exclusion restrictions.⁸ In particular, for each country dyad, we use the log of the distance between capital cities in kilometers and dummies indicating the existence of colonial links after 1945, common official language, and border sharing. All variables are from the GeoDist dataset, built by CEPII⁹.

⁸We follow the literature in making this symmetry assumption. However, we acknowledge the fact that the effect of geographical distance can be asymmetric: going from poor country A to rich country B might be more costly than going from rich country B to poor country A. This might be more problematic for south-north flows, while it should be less important for north-north and south-south corridors.

⁹Available at <http://www.cepii.fr>

3.2 Stylized Facts

Table 1 presents the descriptive statistics. As explained above, our sample includes only half of the net migration flows to avoid counting the same corridor twice. By excluding the main diagonal (where origin and destination coincide), we end up with 18,915 country pairs. About 30 percent of them exhibit positive net flows, and the remaining 70 percent show a zero balance due to the absence of both inflows and outflows. The highest net balances in the data are the corridors from the Philippines to the United States for highly skilled migrants (with a log ratio of 12.7) and from Mexico to the United States for less educated migrants (with a log ratio of 12.5). Each of them roughly represents a net inflow of more than 300,000 people into the United States over the nineties (and conversely a net outflow of 300,000 for the Philippines and Mexico). Table 1 also shows that wage disparities are greater for college graduates than for the less educated. As for the geographical variables, the average geographic distance between two countries is equal to 8.7 thousand kilometers, only one percent of our country pairs have a colonial relationship, 1.5 percent share a common border, and 15 percent share a common official language.

To have better insight into our main variable of interest for the 195 countries in the sample, Figure 1 shows the density of PCA for two groups of countries: high-income OECD countries (referred to as the North) and other countries (referred to as the South).¹⁰ The North exhibits a higher average quality of governance and a much lower dispersion than the South. As far as the dependent variable is concerned, Figure 2 plots the density of the balance for both high-skilled and low-skilled migrants.¹¹ They are both centered around zero and it is easy to note that the dispersion is slightly higher for the low-skilled than for the high-skilled, meaning that the balances for the high-skilled tend to be smaller than those for the less educated. Finally, we combine the information on balances and quality of institution and we plot them separately for the high-skilled (Figure 3) and the low-skilled (Figure 4). Both graphs show a clear positive correlation, suggesting that, for each country pair, the country with the higher quality of institutions tends to experience a surplus of immigrants. The econometric analysis of the next section will also shed light on potential differences between the high-skilled and the low-skilled.

4 Results

This section presents the results of our empirical analysis. Our main variables of interest are the migration balances of college graduates and less educated migrants, measured as logged

¹⁰The classification is available at <http://data.worldbank.org/about/country-classifications/country-and-lending-groups>.

¹¹We use only non-zero values.

net migrant flows ($\ln M_{ji} - \ln M_{ij}$) observed between the years 1990 and 2000. In the first sub-section, we find that migration balances are positively correlated with the six indicators of governance and with the synthetic indicator (PCA) derived from our principal component analysis. We then show that this positive relation is reinforced, in the case of the high-skilled, when an instrumental variable strategy is used, suggesting a causal impact of governance on high-skilled net migration flows. In the second sub-section, we disentangle the effect of home and destination institutional quality on both outflows and inflows.

4.1 Governance and Net Migration Flows

Table 2 presents the OLS estimation results of equation (3) using the six different measures of governance provided in Kaufmann et al. (2009). Panel A gives the results for college graduates and panel B shows the results for the less educated. Our set of controls includes the differences in skill-specific wage rates, populations of stayers left behind, sizes of the network/diaspora, and Mills ratios. Notice that the coefficients for these variables have intuitive signs. Net migration flows increase with the "destination-origin", skill-specific wage gap. The higher the wage in destination i with respect to origin j , the greater the net migration flow from j to i . Net migration flows also increase with the difference in network sizes. If the diaspora from j to i exceeded that of i to j in 1990, this spurred net migration flows from j to i during the nineties (see Carrington, Detragiache, and Vishwanath (1996); Beine et al. (2011)). The estimated coefficients are greater for college graduates than for the less educated. Finally, the difference in Mills ratios (ΔIMR) is always significant in Panel B; it is significant in two regressions of panel A, meaning that it is important to account for sample selection, especially for the low-skilled. Interestingly, the "destination-origin" difference in governance quality between two countries is always associated with greater net migration inflows, whatever the indicator of governance quality. The highest coefficients are obtained with "Government Efficiency", "Control of Corruption", and "Rule of Law". However, since the correlation between the six governance indicators is very large, in column (7), we perform additional OLS regressions using the synthetic governance indicator (PCA). In this case as well, the migration balance is positively associated with the "destination-origin" difference in PCA.

As explained above, OLS regressions might suffer from endogeneity problems. The positive correlation rates between governance and migration balances identified in Table 2 could be driven by a reverse causation link between these two variables, or by the effect of an omitted variable. The first problem may arise if migration affects home and/or foreign institutions (e.g. Li and Hale, 2005; Spilimbergo, 2009; Docquier et al., 2016). The second emerges if any unobserved bilateral component influences both the difference in the quality of institutions and the

net migration flows. The use of net migration flows substantially mitigates this problem, since symmetric unobserved characteristics cancel each other out when taking the difference between inflows and outflows. Still, some asymmetric unobserved characteristics could drive the results. To solve these problems, we use a similar approach to Biavaschi, Giuliotti, and Siddique (2013): we perform a two-stage-least-squares (2SLS) analysis and instrument the distance in the synthetic governance indicator with the distance in the complexity of the destination and origin countries' names.¹² In particular, we measure the complexity of country names by the English Scrabble score of each country's English name and take the difference between origin and destination.¹³ Our identification strategy relies on two key assumptions: the first is that cross-country variation in the quality of institutions is correlated with the complexity of country names (and so the distance in the complexity of country names is correlated with the distance in the quality of institutions). The second is that the complexity of a country's name is not correlated with migration flows.

With respect to the first requirement, different papers argue that language can affect the values at the basis of institutions and their quality. For instance, Chen (2013) shows that linguistic traits affect economic and non-economic attitudes, and Tabellini (2008) finds that language affects the values underlying the general morality, defined as *"the universal applicability of rules of just conduct"*. Following this reasoning, a more complex language might have caused more difficulties in coordination and comprehension, thus limiting the development of institutions. At the same time, a more complex language might have contributed to developing more complex and thus less efficient institutions (Room, 2015). In both cases, the Scrabble score of a country name should be negatively correlated with governance quality. Our analysis supports this hypothesis: Figure 5 shows a negative relationship between the complexity of country names and

¹²Biavaschi et al. (2013) instrument the Americanization of migrants' names using the English Scrabble index of each original migrant name.

¹³More in detail, we assign the English Scrabble score to each letter composing the English name of a country and then add them up. For example, low Scrabble scores are Israel (6), Peru (6), and Austria (7), while high Scrabble scores are Zimbabwe (26), Kyrgyzstan (30), and Mozambique (34). We believe that it would have been more appropriate to use the Scrabble index in the language of the origin country to evaluate the complexity of the name of the destination in the origin country's official language (in other words, we would have used "Francia" - Scrabble score: 12 - for the flow of Italians emigrating to France). However, there are several problems to implement this strategy: i) Scrabble does not exist in all languages (e.g. Chinese); ii) datasets containing the name of countries in different languages have limited language coverage; iii) sometimes, origin countries have more than one official language (e.g. Belgium). Therefore, the English Scrabble and the English name are the easiest solution to avoid making more important assumptions and cutting the sample size of our estimations. This choice is not as extreme as one might think for two reasons: first, the correlation of the Scrabble scores in different languages is more than 95%, which means that the difference in the complexity of a given letter across different languages is quite small; second, excluding some notable exceptions (e.g. Germany, Croatia, etc.), the names of countries tend to be quite consistent across different languages.

the quality of institutions (and this negative correlation is significant at 5%).

With respect to the second condition to have a proper 2SLS, we need the bilateral difference in the complexity of country names to be uncorrelated with migration flows. In Figures 6 and 7, we plot outflows against the complexity of country names, both for high-skilled and low-skilled migrants. No clear pattern can be detected and the correlation is not significantly different from zero at the 5% level. This missing correlation is unlikely to be driven by the effect of migration on the Scrabble, since migrants cannot change the names of countries.¹⁴ Therefore, migrants do not seem to take into account the complexity of country names when choosing where to migrate. One further threat to fulfill this condition is that the correlation between the distance in the Scrabble and the distance in governance quality comes entirely from linguistic proximity. This can be potentially problematic because linguistic proximity is correlated with migration costs and thus with migration flows. Using the linguistic proximity measures provided by Melitz and Toubal (2014), we observe that the correlation with the Scrabble distance of country names is below 15% and never significant. Therefore, our instrument does not seem to capture linguistic proximity¹⁵ and it proves to be significantly correlated with the endogenous variable and uncorrelated with the outcome variable – as the assumptions of 2SLS require.

Column 1 of Table 3 presents the first stage of our 2SLS regression. As expected, the bilateral distance in the complexity of country names is positively and significantly correlated to the distance in the quality of the institutions. In most cases, the first stage F-statistic is in line with the usual thresholds both for the high-skilled and the low-skilled. Looking at the estimates of the second stage in column 1 of Table 4, we observe that an increase in the "destination-origin" difference in governance quality induces a greater high-skilled net migration flow. For the low-skilled instead, this relation remains positive, but it becomes insignificant. The magnitude of the coefficients is greater than in Table 2, suggesting that OLS estimates can be downward biased by a negative reverse effect of migration balances on the quality of institutions. Therefore, the 2SLS procedure confirms the positive causal relationship between the quality of institutions and high-skilled net migration flows, while the estimates for the low-skilled are still positive, but lose significance.

Besides endogeneity, some other issues might bias our results. First, we adjust our in-

¹⁴Moreover, the extent to which human capital *per se* can affect institutions has been found quite limited (Acemoglu, Gallego, and Robinson, 2014)

¹⁵In any case, if the effect of linguistic proximity is symmetric, its effect is canceled out anyway when taking the difference between inflows and outflows. If instead it is asymmetric (as shown for English-speaking destinations by Adsera and Pytlikova (2015)), this could have been more problematic in presence of a strong correlation between the distance in the Scrabble and the linguistic proximity. In the following, we provide evidence that this is not problematic in our case.

flows and outflows to take into account dynamics that may influence the change in the stock of migrants, such as mortality. If there is a correlation between the size of the adjustment and the quality of institutions, the results may be biased. To check for this issue, we re-estimate our model using only the unadjusted (observed) flows. Both for high-skilled and low-skilled, the results remain unchanged (Column 2, Table 4). Second, part of the cells used have been imputed. If the probability of imputing a cell is correlated to the quality of institutions and the instrument, this might cause biased estimates. In column 3 of Table 4, we use only non-imputed cells for our estimations, and we see that, despite the decrease in the number of observations (divided by 10), our results hold. Third, besides the fact that our instrument is not significantly correlated with linguistic proximity, it is worth checking whether our results hold when ruling out the possibility that linguistic proximity exerts a differential effect on migration flows when the destination has English as an official language, as shown by [Adsera and Pytlikova \(2015\)](#). To check this, we exclude from the sample all observations in which either i or j have English as an official language.¹⁶ Column 4 of Table 4 shows that by excluding the possibility of asymmetric effects due to linguistic proximity, the results are still confirmed, despite the decrease by one third of the sample size. Fourth, visa restrictions could represent an important omitted variable for the estimations. While our instrument should be orthogonal to any asymmetry in visa restrictions, it can be interesting to look at the effect of this variable and check whether indeed there is any asymmetric effect. We use the dummy variable constructed by [Neumayer \(2005\)](#) capturing the need of a visa to travel from an origin to a destination. The difference between origin and destination visa policies takes value 1 if there is a restriction in the origin country, -1 if there is a restriction in the destination country, and zero if there are restrictions in both countries or in none. While our results on the quality of institutions still hold, this additional variable is not significant for either the high-skilled or the low-skilled (Column 5 of Table 4).

4.2 Governance and Gross Migration Flows

Having demonstrated that net migration flows are affected by governance quality, we now disentangle the effect of home and foreign institutions on both inflows and outflows. In particular, we estimate equation (3) using the PPML estimation method developed by [Santos Silva and Tenreyro \(2006\)](#).¹⁷ Table 5 (column 1 for high-skilled and column 2 for low-skilled migrants) shows that home institutions are negatively correlated with migration outflows, i.e. more virtu-

¹⁶ Antigua and Barbuda, Australia, the Bahamas, Barbados, Belize, Botswana, Cameroon, Canada, Dominica, Fiji, Gambia, Ghana, Grenada, Guyana, Ireland, Jamaica, Kenya, Lesotho, Liberia, Malawi, Malta, Mauritius, Namibia, New Zealand, Nigeria, Papua New Guinea, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, the Seychelles, Sierra Leone, Singapore, the Solomon Islands, South Africa, Swaziland, Tanzania, Tonga, Trinidad and Tobago, Uganda, the United Kingdom, Vanuatu, Zambia, and Zimbabwe.

¹⁷ Please note that PPML works only with the dependent variable in levels.

ous countries have fewer emigrants. When looking at foreign institutions instead, we observe that high-skilled emigrants tend to choose destinations with higher-quality institutions. Instead, this correlation is not significant for the low-skilled. The results for inflows are complementary to those for outflows (columns 3 and 4, Table 5). College graduates flow into countries with better institutions and avoid those with bad governance quality. The low-skilled instead tend to look only at the quality of institutions in the origin country.

These results suggest that the highly skilled care both about home and foreign institutions, while the low-skilled seem to take into account only the institutional quality in their origin country. This can be the result of the low-skilled having lower access to information on foreign institutions. For example, economic conditions might influence the possibility of acquiring relevant data and low educational attainment may limit the ability to process information and to make an informed decision. This skill difference is also important in our context: for the low-skilled, the difference in the quality of institutions between countries matters less than for the highly skilled and this may be the cause of the insignificant coefficient for the net setting.

Finally, the coefficients for all the other variables have the expected signs. More specifically, the wage effect is negative for outflows, meaning that an increase in wages in the home country will retain both high- and low-skilled migrants. Accordingly, the effect is opposite for inflows, so that higher wages at home increase the number of immigrants. As expected, the diaspora effect is positive and significant for outflows, since the higher the diaspora at destination, the larger the outflow to that country. The same applies to the foreign diaspora: the more foreign citizens live in the home country, the larger the inflow from the foreign country to the home country.

5 Conclusion

Bilateral differences in governance quality impact the size and educational structure of bilateral *net* migration flows. Therefore, the quality of institutions can affect human capital stocks through migration. This paper shows that countries with better governance quality experience positive net flows of college graduate migrants, while countries with worse institutions tend to experience negative net flows. College graduates are more reactive to governance quality and take into account both home and destination governance quality when choosing to emigrate. The low-skilled seem to focus instead only on local conditions in their choices. Therefore, bilateral differences in institutional quality are more important for high-skilled than for low-skilled net migration flows.

References

- Acemoglu, D., F. Gallego, and J. A. Robinson (2014). Institutions, Human Capital and Development. *Annual Reviews of Economics* 6, 875–912.
- Acemoglu, D., S. Johnson, and J. A. Robinson (2005). Institutions as a Fundamental Cause of Long-Run Growth. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth*, Volume 1 of *Handbook of Economic Growth*, Chapter 6, pp. 385–472. Elsevier.
- Acemoglu, D., S. Johnson, J. A. Robinson, and P. Yared (2005). From Education to Democracy? *American Economic Review* 95(2), 44–49.
- Adsera, A. and M. Pytlikova (2015). The Role of Language in Shaping International Migration. *Economic Journal*, forthcoming.
- Artuc, E., F. Docquier, C. Ozden, and C. Parsons (2015). A Global Assessment of Human Capital Mobility: The Role of Non-OECD Destinations. *World Development* 65(C), 6–26.
- Barro, R. J. and J.-W. Lee (2010). A New Data Set of Educational Attainment in the World, 1950–2010. NBER Working Papers 15902, National Bureau of Economic Research, Inc.
- Beine, M., F. Docquier, and C. Ozden, (2011). Diasporas. *Journal of Development Economics* 95(1), 30–41.
- Belot, M. V. K. and T. J. Hatton (2012). Immigrant Selection in the OECD. *Scandinavian Journal of Economics* 114(4), 1105–1128.
- Bertoli, S. and J. Fernández-Huertas Moraga (2013). Multilateral Resistance to Migration. *Journal of Development Economics* 102, 79–100.
- Bertoli, S. and J. Fernández-Huertas Moraga (2015). The Size of the Cliff at the Border. *Regional Science and Urban Economics* 51(C), 1–6.
- Biavaschi, C., C. Giulietti, and Z. Siddique (2013, November). The Economic Payoff of Name Americanization. IZA Discussion Papers 7725, Institute for the Study of Labor (IZA).
- Carrington, W. J., E. Detragiache, and T. Vishwanath (1996). Migration with Endogenous Moving Costs. *American Economic Review* 86(4), 909–30.
- Chen, M. K. (2013). The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets. *American Economic Review* 103(2), 690–731.
- Coulombe, S. and J.-F. Tremblay (2009). Migration and Skills Disparities across the Canadian Provinces. *Regional Studies* 43(1), 5–18.
- Docquier, F., E. Lodigiani, H. Rapoport, and M. Schiff (2016). Emigration and Democracy. *Journal of Development Economics*, forthcoming.
- Docquier, F., O. Lohest, and A. Marfouk (2007). Brain Drain in Developing Countries. *World Bank Economic Review* 21(2), 193–218.

- Docquier, F., B. L. Lowell, and A. Marfouk (2009). A Gendered Assessment of Highly Skilled Emigration. *Population and Development Review* 35(2), 297–321.
- Docquier, F. and H. Rapoport (2003). Ethnic Discrimination and the Migration of Skilled Labor. *Journal of Development Economics* 70(1), 159–172.
- Grogger, J. and G. H. Hanson (2011). Income Maximization and the Selection and Sorting of International Migrants. *Journal of Development Economics* 95(1), 42–57.
- Hall, R. E. and C. I. Jones (1999). Why Do Some Countries Produce So Much More Output Per Worker Than Others? *The Quarterly Journal of Economics* 114(1), 83–116.
- Helpman, E., M. Melitz, and Y. Rubinstein (2008). Estimating Trade Flows: Trading Partners and Trading Volumes. *The Quarterly Journal of Economics* 123(2), 441–487.
- Hendricks, L. (2004). A Database of Mincerian Earnings Regressions. Technical report, www.lhendricks.org/Mincer.htm.
- Hirschman, A. (1970). *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press.
- Kaufmann, D., A. Kraay, and M. Mastruzzi (2009). Governance Matters VIII: Aggregate and Individual Governance Indicators 1996-2008. Policy Research Working Paper Series 4978, The World Bank.
- Kaufmann, D., A. Kraay, and P. Zoido-Lobaton (1999). Governance Matters. Policy Research Working Paper Series 2196, The World Bank.
- Li, X. and J. M. Hale (2005). Does Brain Drain Lead to Institutional Gain? A Cross-Country Empirical Investigation. Mimeo, Queens University.
- Mattoo, A., I. C. Neagu, and Â. Özden (2008). Brain Waste? Educated Immigrants in the US Labor Market. *Journal of Development Economics* 87(2), 255–269.
- Mayda, A. (2010). International Migration: a Panel Data Analysis of the Determinants of Bilateral Flows. *Journal of Population Economics* 23(4), 1249–1274.
- McFadden, D. (1984). Econometric Analysis of Qualitative Response Models. In Z. Griliches and M. Intriligator (Eds.), *Handbook of Econometrics*, Volume 2. Amsterdam: Elsevier.
- McKenzie, D. (2007). Paper Walls Are Easier to Tear Down: Passport Costs and Legal Barriers to Emigration. *World Development* 35(11), 2026–2039.
- Melitz, J. and F. Toubal (2014). Native Language, Spoken Language, Translation and Trade. *Journal of International Economics* 93(2), 351–363.
- Neumayer, E. (2005). Visa Restrictions and Bilateral Travel. *The Professional Geographer* 62(2), 1–11.
- Pedersen, P. J., M. Pytlikova, and N. Smith (2008). Selection and Network Effects—Migration Flows into OECD Countries 1990-2000. *European Economic Review* 52(7), 1160–1186.

- Room, G. (2015). Complexity, power and policy. In R. Geyer and P. Cairney (Eds.), *Handbook on Complexity and Public Policy*, pp. 19–31. Cheltenham: Elgar.
- Santos Silva, J. M. C. and S. Tenreiro (2006). The Log of Gravity. *The Review of Economics and Statistics* 88(4), 641–658.
- Shleifer, A., F. L. de Silanes, and R. La Porta (2008). The Economic Consequences of Legal Origins. *Journal of Economic Literature* 46(2), 285–332.
- Spilimbergo, A. (2009). Democracy and Foreign Education. *American Economic Review* 99(1), 528–43.
- Tabellini, G. (2008). Institutions and Culture. *Journal of the European Economic Association Papers and Proceedings*.

Figure 1: Quality of Institutions

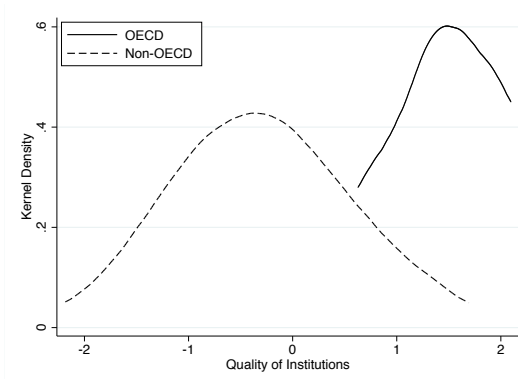


Figure 2: Balance

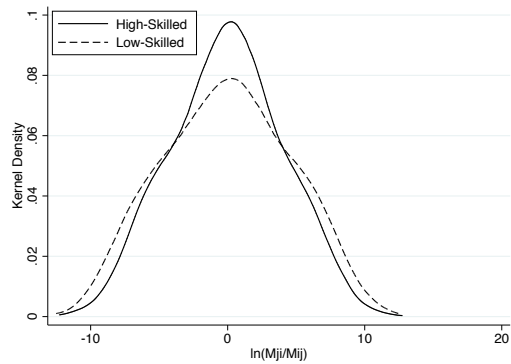


Figure 3: H-S Balance and ΔPCA

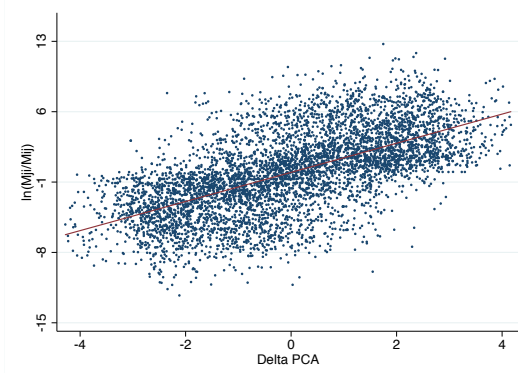


Figure 4: L-S Balance and ΔPCA

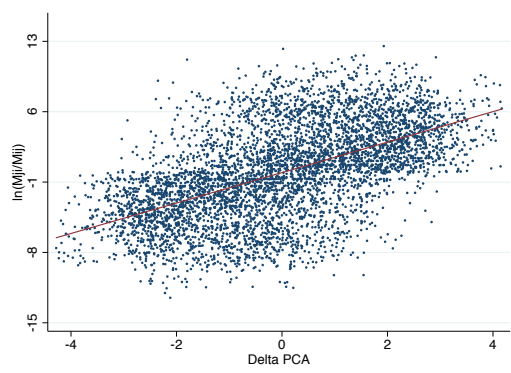


Figure 5: PCA and the Complexity of Country Names

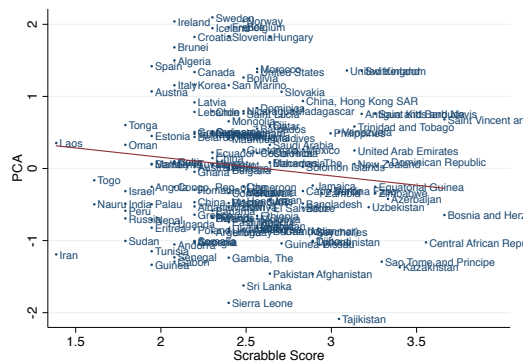


Figure 6: High-Skilled Outflow and PCA

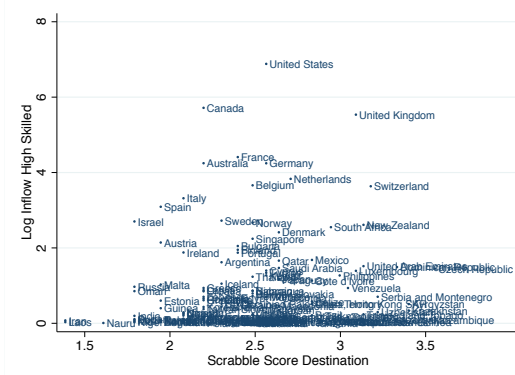


Figure 7: Low-Skilled Outflow and PCA

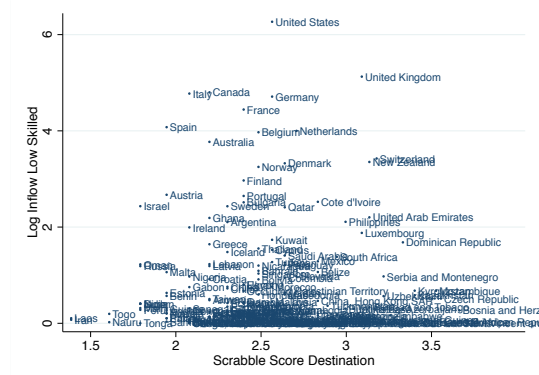


Table 1: Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Net high-skilled	18,915	0.0108	2.1790	-12.7328	11.6493
Net high-skilled Down.	18,915	0.0091	2.0304	-12.2595	11.1927
Net low-skilled	18,915	0.0033	2.4537	-12.5255	12.2559
Net low-skilled Down.	18,915	0.0112	2.5422	-12.7413	12.2585
ΔWage_{hs}	18,915	0.0332	41.1990	-206.1103	204.7582
ΔWage_{ls}	18,915	-0.0042	16.5172	-63.9722	63.7881
$\Delta \text{Log Natives}_{hs}$	18,915	0.0440	4.8240	-17.0866	17.0866
$\Delta \text{Log Natives}_{ls}$	18,915	0.0152	5.7443	-19.1104	19.1104
$\Delta \text{Log Diaspora}$	18,915	0.0151	2.1061	-12.3563	10.8819
ΔPCA	16,836	-0.0086	1.4180	-4.2308	4.2800
$\Delta \text{Control of Corruption}$	17,955	-0.0052	1.3962	-4.2007	4.1665
$\Delta \text{Rule of Law}$	18,721	-0.0057	1.3919	-4.1124	4.1081
$\Delta \text{Political Stability}$	16,836	-0.0095	1.3922	-4.1294	4.0502
$\Delta \text{Violence \& Accountability}$	18,721	-0.0105	1.3957	-3.6341	3.6386
$\Delta \text{Government Efficiency}$	17,955	-0.0032	1.3949	-4.1646	4.1758
$\Delta \text{Regulatory Quality}$	17,955	-0.0061	1.3839	-4.3315	4.5501
Log Distance	17,766	8.7660	0.7769	4.0879	9.9010
Colonial Links	17,766	0.0109	0.1039	0	1
Common Language	17,766	0.1548	0.3617	0	1
Common Border	17,766	0.0157	0.1245	0	1

Notes: This table presents descriptive statistics for the variables used in the analysis.

Table 2: Net Migration Flows and Quality of Institutions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
PANEL A: High-skilled							
Δ Control of Corruption	0.257*** (0.013)						
Δ Rule of Law		0.208*** (0.012)					
Δ Pol. Stability			0.141*** (0.012)				
Δ Voice and Account.				0.166*** (0.010)			
Δ Governm. Effectiveness					0.271*** (0.013)		
Δ Regul. Quality						0.225*** (0.012)	
Δ PCA							0.254*** (0.013)
Δ Wage	0.002*** (0.000)	0.003*** (0.000)	0.003*** (0.001)	0.005*** (0.001)	0.002*** (0.000)	0.002*** (0.000)	0.003*** (0.001)
Δ Population	-0.008*** (0.003)	-0.000 (0.003)	0.002 (0.003)	0.002 (0.003)	-0.013*** (0.003)	-0.011*** (0.003)	-0.006* (0.003)
Δ Diaspora	0.554*** (0.012)	0.570*** (0.012)	0.585*** (0.012)	0.575*** (0.012)	0.560*** (0.012)	0.573*** (0.012)	0.562*** (0.012)
ΔIMR	0.046 (0.092)	0.066 (0.096)	-0.233** (0.106)	0.112 (0.101)	0.123 (0.094)	0.037 (0.094)	0.200* (0.104)
R ²	0.46	0.45	0.45	0.45	0.46	0.46	0.46
Observations	17,020	17,578	15,931	17,578	17,020	17,020	15,931
PANEL B: low-skilled							
Δ Control of Corruption	0.301*** (0.018)						
Δ Rule of Law		0.256*** (0.016)					
Δ Pol. Stability			0.165*** (0.016)				
Δ Voice and Account.				0.241*** (0.013)			
Δ Governm. Effectiveness					0.338*** (0.019)		
Δ Regul. Quality						0.236*** (0.017)	
Δ PCA							0.322*** (0.018)
Δ Wage	0.030*** (0.002)	0.033*** (0.002)	0.039*** (0.002)	0.040*** (0.002)	0.032*** (0.002)	0.037*** (0.002)	0.036*** (0.002)
Δ Population	-0.019*** (0.003)	-0.014*** (0.003)	-0.009** (0.003)	-0.011*** (0.003)	-0.021*** (0.003)	-0.020*** (0.003)	-0.017*** (0.003)
Δ Diaspora	0.464*** (0.015)	0.487*** (0.015)	0.466*** (0.015)	0.498*** (0.015)	0.496*** (0.015)	0.492*** (0.015)	0.498*** (0.015)
ΔIMR	0.918*** (0.140)	1.059*** (0.144)	0.797*** (0.153)	1.277*** (0.146)	1.301*** (0.148)	1.152*** (0.151)	1.397*** (0.156)
R ²	0.35	0.34	0.34	0.34	0.35	0.34	0.35
Observations	17,020	17,578	15,931	17,578	17,020	17,020	15,931

Notes: All regressions are run by OLS. The dependent variable is the log net migration flow. Panel A shows results for high-skilled migrants and Panel B shows results for low-skilled migrants. We use the 6 different measures of governance from Kaufmann et al. (1999) and the principal component derived from them. For more details on the governance indicators and controls, see Section 3. Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 3: IV: First Stage

	(1)	(2)	(3)	(4)	(5)
	Adjusted Migrants	Observed Migrants	Non-Imputed Cells	Non-English Dest. Orig.	Visa Restrictions
PANEL A: High-skilled					
Δ Scrabble	0.052*** (0.009)	0.052*** (0.009)	0.117** (0.047)	0.030** (0.016)	0.093*** (0.013)
Δ Wage	0.003*** (0.001)	0.003** (0.001)	0.021*** (0.001)	0.003*** (0.000)	0.006** (0.002)
Δ Population	0.004** (0.002)	-0.004*** (0.002)	-0.044*** (0.009)	-0.010*** (0.002)	0.006*** (0.001)
Δ Diaspora	0.152*** (0.004)	0.152*** (0.004)	0.104*** (0.008)	0.138*** (0.005)	0.138*** (0.004)
Δ VISA					0.114*** (0.019)
ΔIMR	-3.670*** (0.059)	-3.670*** (0.059)	-7.340*** (0.305)	-3.397*** (0.075)	-3.842*** (0.058)
R^2	0.47	0.47	0.39	0.43	0.49
F-Statistic	16.01	15.85	6.16	12.29	20.63
Observations	15,931	15,931	1,830	10,273	15,225
PANEL B: low-skilled					
Δ Scrabble	0.024** (0.012)	0.024** (0.007)	0.291*** (0.039)	0.032** (0.014)	0.055*** (0.012)
Δ Wage	0.015*** (0.001)	0.015*** (0.001)	0.001 (0.002)	0.018*** (0.001)	0.013*** (0.001)
Δ Population	0.012*** (0.001)	0.012*** (0.001)	0.041*** (0.004)	0.016*** (0.002)	0.017*** (0.001)
Δ Diaspora	-0.204*** (0.006)	-0.204*** (0.006)	-0.385*** (0.020)	-0.186*** (0.007)	-0.225*** (0.005)
Δ VISA					-0.058*** (0.017)
ΔIMR	-3.524*** (0.068)	-3.524*** (0.068)	-5.782*** (0.265)	-3.127** (0.083)	-3.738*** (0.068)
R^2	0.59	0.59	0.58	0.57	0.61
F-Statistic	15.86	4.21	56.53	5.18	22.14
Observations	15,931	15,931	1,830	10,273	15,225

Notes: This table shows the results of the first stage of the 2SLS strategy presented in Section 4, Panel A for high-skilled migrants and Panel B for low-skilled migrants. Our instrument for the distance in the quality of institutions is the distance in the Scrabble index of the name of the origin and destination countries. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: IV: Scrabble distance in country names

	(1)	(2)	(3)	(4)	(5)
	Adjusted Migrants	Observed Migrants	Non-Imputed Cells	Non-English Dest.Orig.	Visa Restrictions
PANEL A: High-skilled					
Δ PCA	2.201*** (0.619)	2.492*** (0.683)	1.833* (1.063)	4.961* (2.746)	1.321*** (0.264)
Δ Wage	0.010*** (0.002)	0.012*** (0.003)	0.035 (0.023)	0.018** (0.009)	0.007*** (0.001)
Δ Population	-0.012** (0.005)	-0.011* (0.006)	0.034 (0.053)	-0.061** (0.030)	-0.011*** (0.004)
Δ Diaspora	0.261*** (0.096)	0.264** (0.106)	0.224** (0.112)	-0.122 (0.383)	0.405*** (0.039)
ΔIMR	7.343*** (2.281)	8.314*** (2.517)	9.235 (7.868)	16.022* (9.333)	4.277*** (1.022)
Δ Visa					-0.022 (0.049)
R^2	0.44	0.45	0.12	0.40	0.22
Observations	15,931	15,931	1,830	10,273	15,225
PANEL B: low-skilled					
Δ PCA	2.597 (1.605)	1.769 (1.368)	3.832*** (0.708)	-1.750 (1.467)	1.056* (0.543)
Δ Wage	0.003 (0.023)	0.011 (0.020)	0.084*** (0.011)	0.076*** (0.027)	0.028*** (0.007)
Δ Population	-0.043** (0.019)	-0.035** (0.016)	-0.159*** (0.031)	0.016 (0.024)	-0.028*** (0.010)
Δ Diaspora	0.960*** (0.327)	0.757*** (0.279)	1.727*** (0.299)	0.091 (0.275)	0.659*** (0.123)
ΔIMR	9.404* (5.665)	6.460 (4.827)	23.339*** (4.376)	-5.179 (4.601)	4.197** (2.038)
Δ Visa					0.075 (0.054)
R^2	0.30	0.31	0.26	0.33	0.30
Observations	15,931	15,931	1,830	10,273	15,225

Notes: This table shows the results of the second stage of the 2SLS strategy presented in Section 4. The dependent variable is the log net migration flow. Panel A shows results for high-skilled migrants and Panel B shows results for low-skilled migrants. Our instrument for the distance in the quality of institutions is the distance in the Scrabble index of the name of the origin and destination countries. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Outflows and Inflows Regressions

	Outflows		Inflows	
	(1)	(2)	(3)	(4)
	high-skilled	low-skilled	high-skilled	low-skilled
PANEL A: high-skilled				
PCA Home	-0.236*** (0.074)	-0.354*** (0.085)	0.958*** (0.115)	0.010 (0.150)
PCA Foreign	0.347*** (0.070)	0.073 (0.105)	-0.343*** (0.097)	-0.246** (0.113)
Δ Wage	-0.011*** (0.002)	-0.016* (0.009)	0.019*** (0.004)	0.090*** (0.011)
Diaspora	0.628*** (0.041)	0.585*** (0.041)	0.215*** (0.052)	0.271*** (0.042)
Colonial Links	0.305* (0.157)	-0.199 (0.348)	0.604*** (0.234)	-0.158 (0.438)
Distance	-0.110 (0.086)	-0.036 (0.144)	-0.136 (0.199)	-0.230 (0.259)
Common Language	0.378** (0.153)	0.054 (0.181)	0.846*** (0.234)	0.217 (0.253)
Common Border	-0.141 (0.245)	0.199 (0.275)	1.149* (0.589)	2.263** (0.908)
Origin Residents	0.202*** (0.033)	0.036 (0.026)	0.447*** (0.073)	0.412*** (0.082)
Destination Residents	0.116** (0.047)	0.024 (0.054)	0.337*** (0.101)	0.158 (0.115)
R ²	0.76	0.19	0.43	0.67
Observations	15,931	15,930	15,931	15,930

Notes: All regressions are run using PPML. The dependent variable is the outflow between two countries in levels. Panel A shows results for high-skilled migrants and Panel B shows results for low-skilled migrants. * p<0.1, ** p<0.05, *** p<0.01.