C E P R E M A P

CENTRE POUR LA RECHERCHE ECONOMIQUE ET SES APPLICATIONS

Document de travail (Docweb) nº 1406

Emigration and democracy

Frédéric Docquier

Elisabetta Lodigiani

Hillel Rapoport

Maurice Schiff

Février 2014

Abstract: International migration is an important determinant of institutions, not considered so far in the empirical growth literature. Using cross-section and panel analysis for a large sample of developing countries, we find that openness to emigration (as measured by the general emigration rate) has a positive effect on home-country institutional development (as measured by standard democracy indices). The results are robust to a wide range of specifications and estimation methods. Remarkably, the cross-sectional estimates are fully in line with the implied long-run relationship from dynamic panel regressions.

Résumé: Les migrations internationales sont un déterminant important des institutions, non encore considéré dans la littérature empirique sur la croissance. Nous procédons à des analyses en coupe transversale ainsi qu'en panel pour un large échantillon de pays en développement et montrons que l'ouverture à l'émigration (mesurée par le taux général d'émigration) a un effet positif sur le développement institutionnel (mesuré par des indicateurs standards de niveau de démocratie). Les résultats sont robustes pour un grand nombre de spécifications et de méthodes d'estimation. De façon remarquable, les estimations cross-sectionnelles sont totalement en phase avec la relation de long terme qui ressort des régressions en panel dynamique.

Emigration and democracy*

Frédéric Docquier[a], Elisabetta Lodigiani[b],
Hillel Rapoport[c] and Maurice Schiff[d]

[a] FNRS and IRES, Université Catholique de Louvain

[b] Ca' Foscari University of Venice and Centro Studi Luca d'Agliano

[c] Paris School of Economics, University Paris 1 Pantheon-Sorbonne and Bar-Ilan University

[d] World Bank; University of Chile and IZA

Revised Version – December 30th, 2013

Abstract

International migration is an important determinant of institutions, not considered so far in the empirical growth literature. Using cross-section and panel analysis for a large sample of developing countries, we find that openness to emigration (as measured by the general emigration rate) has a positive effect on home-country institutional development (as measured by standard democracy indices). The results are robust to a wide range of specifications and estimation methods. Remarkably, the cross-sectional estimates are fully in line with the implied long-run relationship from dynamic panel regressions.

JEL codes: O15, O43, F22.

Keywords: Migration, Institutions, Democracy, Development.

*Corresponding author: Hillel Rapoport, Paris School of Economics: hillel.rapoport@psemail.eu.
An early version of this paper circualted as World Bank Policy Research Paper No 5557 (Docquier et al., 2013); it is part of the World Bank Research Program on International Migration and Development. We thank the editor, three anonymous referees, and Michel Beine, Eckhardt Bode, Emanuele Forlani, Jeffrey Frieden, David McKenzie, Anna Maria Mayda, Caglar Ozden, Robert Vermeulen, Jeffrey Williamson, and participants at seminars and conferences at the World Bank, Louvain, Luxembourg, Paris 1, Maastricht, Boston University, Kiel, Georgetown, Harvard and Venice for comments and suggestions. We are grateful to Pierre Yared and to Herbert Bruecker, Stella Capuano and Abdeslam Marfouk for sharing their data with us.

1 Introduction

Recent research has emphasized the importance of institutions for economic growth (Acemoglu, Johnson and Robinson, 2005, Rodrik, 2007) and explored the determinants of institutions. This paper argues that migration is an important determinant of institutions, not considered so far in the empirical growth literature.

Migration first affects institutions by providing people with exit options, thereby reducing their incentives to voice (Hirschman, 1970); for those who stay, the incentives to voice are also reduced by the possibility of receiving remittance income (which can act as a safety net), further alleviating social, political and economic pressures to reform. For example, it is commonly argued that emigration to the United States contributed to delay political change in countries such as Mexico or Haiti.[1] On the other hand, once abroad, migrants can engage in political activities (e.g., lobby the host-country government to encourage or block financial aid, or impose economic sanctions) which affect the institutional evolution of their home country, for good or bad.[2] International migration also creates fiscal competition across national jurisdictions, possibly limiting the scope for rent-seeking and, as such, contributing to better institutional and governance outcomes.[3]

A second channel through which migration affects institutions has to do with the fact that migration is a selective process. Migrants are not randomly drawn out of a country's population but tend to self-select along a variety of dimensions. First, migrants may be politically self-selected; in the context of developing and transition

[1] See for example Hansen (1988) on Mexico and Fergusson (2003) on Haiti.

[2] A well-known example is the very active anti-Castro lobby in the United States which has succeeded in maintaining a total embargo on economic relations with Cuba (Eckstein, 2009; Haney and Vanderbush, 1999, 2005; Vanderbush, 2009). A lesser known example is the role of the Croatian diaspora in the US and the EU, which strongly supported secession from the former Yugoslavia and the transition to a market-led economy, provided huge financial support to Tudjman's Croatian Democratic Union (CDU) party and, following the latter's victory in the first post-communist elections in 1990, saw its efforts rewarded by the allocation of 12 out of 120 seats at the national assembly to diaspora Croats (Djuric, 2003; Ragazzi, 2009). Diasporas may also at times side with a specific group in a civil war. For instance, Irish Catholics in the United States have historically provided financial and other forms of support to the Catholic community in Northern Ireland, encouraging the most radical factions and therefore making it more difficult to reach a peace agreement (Holland, 1999; Wilson, 1995). Similar analyses have been proposed notably for Lebanon and Sri Lanka.

[3] The idea of migration as a personal response to political and economic repression has a long tradition in economics and political science (see Vaubel, 2008). Recent political economy models of the interaction between emigration, institutions and development include Esptein et al. (1999), Docquier and Rapoport (2003), Mariani (2007) and Wilson (2011).

countries they tend to positively self-selected in terms of preferences for democracy (e.g., Hirschman, 1993). Second, migrants are typically positively self-selected on education. Given that more educated individuals – and the middle class in general (Easterly, 2001) – tend to have a higher degree of political participation and to contribute a greater deal to public policy debates, emigration is likely to hurt the quality of domestic institutions as well as the process through which good policies are formulated and implemented. On the other hand, migration raises the expected return to human capital, thus inducing people to invest more (or more people to invest) in education (e.g., Mountford, 1997, Beine et al., 2001, 2008, Katz and Rapoport, 2005) and to reallocate talent toward productive and internationally transferable skills (Mariani, 2007); such effects on the skill distribution can mitigate or even reverse any adverse brain drain impact on political institutions. Third, another characteristic on which migrants are not randomly self-selected is ethnicity, with an over-representation of ethnic minorities among emigrants. This tends to recompose the home-country population towards more homogeneity, again, for good or bad.[4]

Finally, emigration increases the home country's exposure to new political values and practices, be it directly, through contacts with return migrants and relatives abroad, or indirectly, through the broader scope of diaspora networks. Such networks have been shown to foster bilateral trade, investment and knowledge flows (see Docquier and Rapoport, 2012, for a review of this literature) and to contribute to the transfer of fertility norms (Fargues, 2007, Beine, Docquier and Schiff, 2013, Bertoli and Marchetta, 2013) and to the diffusion of preferences for democracy.

In particular, two recent micro studies find supportive evidence of a democracy-diffusion effect of emigration. In the context of Cape Verde,[5] Batista and Vicente (2011) took advantage of a survey on perceived corruption in public services to set up a "voting experiment": respondents to the survey were asked to mail a pre-stamped

[4]In the penultimate paragraph of their article on "artificial states", Alesina, Easterly and Matuszeski (2011) write: "probably the single most important issue that we have not addressed is that of migrations. One consequence of artificial borders is that people may want to move, if they can. ... In some cases, migrations that respond to artificial borders may be partly responsible for economic costs, wars, dislocation of people, refugee crises and a host of undesirable circumstances. ... But sometimes the movement of people may correct for the artificial nature of borders."

[5]Cape Verde is a nine-island tropical country off the coast of West Africa with a population of half a million, good institutional scores by African standards, and a long tradition of migration. Current migrants represent one-fifth of the population, and brain drain rates are extremely high – 67% in Docquier and Marfouk (2006) – and remain very high (60%) even after excluding people who emigrated before age 18 and acquired their tertiary education abroad (Beine et al., 2007).

postcard if they wanted the results of the survey to be made publicly available in the national media. Controlling for individual, household and locality characteristics, Batista and Vicente (2011) regressed participation in the voting experiment – which they interpret as demand for accountability – on migration prevalence at the locality level. They show that current as well as return migrants significantly increase participation rates, and more so for the latter. Interestingly, they find that only migrants to the US seem to make an impact, while migrants to Portugal, the other main destination, do not. The other context we report on is that of Moldova, a former Soviet Republic with virtually no emigration before 1990 which has seen a recent surge in migration outflows, estimated at half-a-million for a population of 3.6 million in 2008. The evidence we present for Moldova comes from the analysis of election outcomes in Omar Mahmoud et al. (2013). They take advantage of the quasi-experimental context in which the episode of massive emigration they analyze took place and of the fact that Moldovan emigration was directed both to the more democratic European Union and to less democratic Russia, allowing for estimating destination-specific effects. They find that past emigration to the West translates into significantly lower share of votes for the communist party at the community level and provide suggestive evidence of information and cultural transmission channels.

The closest related paper is Spilimbergo (2009), who also adopts a cross-country approach and shows that foreign-trained individuals promote democracy at home if foreign education was acquired in democratic countries. While he does not identify the mechanisms at work, he suggests a number of possible channels (e.g., access to foreign media, acquisition of norms and values while abroad that diffuse at home upon return, willingness to preserve the quality of one's network abroad, etc.) that can be generalized to other migration experiences as well. Our paper is similar in spirit and execution, with important conceptual differences. First, we estimate the effect of emigration on home-country institutions for all migrants, not just foreign students, meaning that we proceed to a larger scale exercise. Second, Spilimbergo's data contains information on the number of people with foreign training living either abroad or in the home country, making it impossible to know whether the effect is due to those staying abroad or to those who returned. In contrast, our emigration variable consists of the lagged accumulated stock of individuals (aged 25+) born in the home country and living abroad, suggesting that the effect of emigration on democracy needs not be driven by return migration. Third, identification in Spilimbergo's

paper fully relies on heterogeneous (or destination-specific) effects. Given that our data set is restricted to OECD destinations which are very homogenous in terms of democratic performance,[6] we cannot test for the effect of emigration to democratic v. non-democratic countries. Our identification strategy relies instead on instrumental variables techniques, as detailed in the methodology section below. Fourth, Spilimbergo finds consistent results only for the "democratic norm at destination" variable, a weighted average of democratic scores at destination which captures whether emigration is directed toward more or less democratic countries. In all his specifications but one, the interaction term between the number of students abroad and the "democratic norm" is not significant. In contrast, our main results are for the volume of migration, suggesting that whether a country has one or twenty percent emigration rate makes a difference, not just whether its emigration is directed toward destinations with higher or lower democracy scores. Incidentally but quite importantly, this also allows us to interpret the magnitude of the estimated effects.

As in Spilimbergo (2009), our methods allow, and indeed force us to examine the overall impact of emigration on home-country institutions. This is composed of the direct and indirect effects detailed in the first paragraphs of this introduction. Section 2 presents the empirical model, discusses the main challenges for the empirical analysis, and describes the data. Section 3 presents the results. Section 4 concludes.

2 Empirical strategy

Our goal is to empirically investigate the effect of emigration on the quality of institutions in the sending country. We will use several indicators of institutional quality, I_t, and measures of openness to emigration, m_t, available for origin country $i = 1, ..., N$ and year $t = 1, ..., T$. In our benchmark regressions, the emigration rate is computed as the sum of emigrants from country i to OECD destination countries j at time t, $\sum_j M_{ij,t}$, divided by the native population of country i, $N_{i,t}$. In this section we present our empirical model, discuss a number of econometric issues and describe the data sources used for the empirical analysis.

[6] For example, only one country (Chile) in our sample of 20 OECD destinations was classified as a "flawed" (as opposed to "true") democracy according to The Economist Intelligence Unit in 2008.

2.1 Model

Our empirical model features the quality of institutions as the dependent variable. It is well known that institutional indicators exhibit some inertia. We need a dynamic regression model to explain their evolution. We augment the dynamic specification used in previous studies (Acemoglu et al., 2005, Bobba and Coviello, 2007, Castello-Climent, 2008, and Spilimbergo 2009) by adding the emigration rate as RHS variable:

$$I_{i,t} = \alpha_0 + \beta_0 I_{i,t-1} + \beta_1 m_{i,t-1} + \beta_2 h_{i,t-1} + \beta_3 X_{i,t-1} + \varepsilon_{i,t} \tag{1}$$

where α_0 is a constant, $h_{i,t}$ is the stock of human capital (measured as the proportion of workers with college education), and $X_{i,t-1}$ and is a vector of time-varying explanatory variables. The lagged dependent variable enters the set of explanatory variables with coefficient β_0 to account for persistence in institutional quality. All explanatory variables are lagged by one period (one period represents five years). Our coefficient of interest is β_1; it captures the effect of the emigration rate on institutional quality at home. The coefficient β_2 captures the effect of human capital on democracy; and β_3 is a vector of parameters associated with a set of additional controls such as GDP per capita, trade openness and ODA flows as share of GDP. Coefficient β_k captures the short-run effect of explanatory variable k on institutions.

Such a dynamic model has been extensively used to explain the dynamics of persistent variables such as the stock of human/physical capital or GDP per capita. If explanatory variables are constant ($m_{i,t} = m_{i,ss}, h_{i,t} = h_{i,ss}$ and $X_{i,t} = X_{i,ss}$ $\forall t$, where subscript ss stands for steady state) and if the coefficient $\beta_0 \in [0; 1]$, then the level of the dependent variable converges towards a long-run or steady state level

$$I_{i,ss} = \frac{\alpha_0 + \beta_1 m_{i,ss} + \beta_2 h_{i,ss} + \beta_3 X_{i,s}}{1 - \beta_0}, \tag{2}$$

which characterizes the long-run relationship between institutions and the RHS variables. Hence, $\beta_k/(1 - \beta_0)$ captures the long-run effect of explanatory variable k.

Estimating (1) requires panel data while estimating (2) can be done in a cross-sectional setting with one observation per country.

2.2 Econometric issues

We will first estimate (2) and (1) using OLS or pooled OLS regressions. However, such regressions raise a number of econometric issues (notably: reverse causality, endogeneity of other regressors, and omitted variables) that might generate inconsistent OLS estimates. These issues and the way we deal with them are discussed below. More generally, cross sectional and panel data techniques have their pros and cons. In a cross-country framework, the underlying steady-state assumptions, albeit questionable, allows to circumvent the difficulty inherent to the endogeneity of the lagged dependent; however, in such a framework the omitted variable issue is severe. In a panel framework, on the other hand, we can better deal with unobserved heterogeneity and characterize the transitional dynamics of institutional quality; however we need to find exogenous instruments that are both country- and time-specific.

Reverse causality. A key issue when using cross-sectional or pooled OLS regressions in our context is the endogeneity of our main variable of interest, the emigration rate. The quality of institutions is likely to affect the desire to emigrate (as most people prefer to live in countries with good institutions) and the possibility to emigrate (as bad institutions or low government effectiveness can be responsible for large administrative costs).[7] This means that a positive or negative correlation between emigration and institutional quality can be driven by reverse causality. Solving this endogeneity issue requires (i) using a two-stage least squares (2SLS) estimation strategy, and (ii) finding a suitable instrument for migration in the first stage.

The philosophy of our 2SLS strategy is the following. In the cross-sectional setting, we focus on the year 2005 and follow Frankel and Romer (1999) to construct a geography-based prediction of bilateral migration stocks, $\widehat{M}_{ij,05}$. The predicted emigration rate, $\widehat{m}_{i,05}$, is then obtained by aggregating bilateral migration stocks over destinations, $\sum_j \widehat{M}_{ij,05}$, and dividing the sum by the native population size in 2005. We use the geography-based predicted rate to instrument $m_{i,05}$ in our first stage regression. This method is now standard in the migration literature (e.g., Beine et al. 2013, Ortega and Peri 2013, Alesina, Harnoss and Rapoport, 2013) and follows a long tradition of predicting trade openness out of bilateral trade flows. Following Rodrik et al. (2004), however, we also include "absolute geography" in our regressions.

In the cross-section settting, the geography-based predictions of bilateral migra-

[7] Fitzgerald et al. (2013) study the political pull factors of international migration in a gravity framework.

tion stocks are obtained from the following pseudo-gravity model:

$$\ln M_{ij,05} = a_0 + a_j + b_1 Lin_{ij} + b_2 Guest_{ij} + b_3 \ln D_{ij} + b_4 \ln P_{i,05} + \epsilon_{ij,05}$$

where Lin_{ij} is a dummy variable equal to 1 if a language is spoken by at least 9% of the population in both countries, $Guest_{ij}$ is a dummy variable equal to 1 if a guest-worker program after 1945 and before the 1980s was observed, $\ln D_{ij}$ is the log of the weighted distance that is equal to the distance between i and j based on bilateral distances between the biggest cities of the two countries (with those inter-city distances being weighted by the share of the city in the total population of the country, see Head and Mayer (2002)), $\ln P_{i,05}$ represents the (log) of the total population at origin in 2005, and a_j is a destination-country fixed effect. Our model does not include origin-country fixed effects because the latter are likely to capture the effect of institutions on emigration decisions.

The presence of a large number of zeroes in bilateral migration stocks gives rise to econometric concerns about possible inconsistent OLS estimates. The most appropriate method to estimate the above model is the Poisson regression by pseudo-maximum likelihood (PPML). We will use the PPML command in Stata which uses the method of Santos Silva and Tenreyro (2011) to identify and drop regressors that may cause the non-existence of the (pseudo-) maximum likelihood estimates. Standard errors are robust and clustered by country pairs. The limitation of this instrumentation strategy is that most of our determinants of bilateral migration stocks are time-invariant. In a panel setting, therefore, we follow Feyrer (2009) and use time fixed effects and interaction between geographic distance and time dummies.

Omitted variables. Estimating (2) and (1) requires defining a set of explanatory variables affecting the quality of institutions. Many explanatory variables have been used in the literature on the determinants of institutions such as GDP per capita, human capital, legal origin dummies, religious variables, latitude, fractionalization indices, etc. A key issue when adding explanatory variables is that they exhibit collinearity (see Alesina et al., 2003).[8] For example, GDP per capita and human capital are collinear, and latitude is correlated with legal origin and fractionalization. Introducing correlated controls can therefore generate identification problems among

[8]For example, Alesina et al. (2003) point out that their index of ethnic fractionalization is highly correlated with latitude and with the log of gdp per capita (which, in addition, is very likely endogenous). Moreover, legal origin dummies are highly correlated with religious variables, etc.

the correlated variables. In a panel setting, we can solve this problem by controlling for time fixed effects, α_t, and country fixed effects, α_i. Although they cannot capture determinants that are both country- and time-specific, they account for may unobservable characteristics that jointly affect emigration and institutions.

In our estimation strategy, we do not consider a within transformation to control for unobserved heterogeneity as results will become far too imprecise for several reasons. First, we know that in a dynamic panel data model, the standard fixed effect estimator is biased and inconsistent in panels with a short time dimension (the so called Nickell bias (Nickell, 1981)). Second, as Hauk and Wacziarg (2009) point out, the within estimator tends to exacerbate the measurement error bias and to understate the impact of explanatory variables in dynamic panel data models with regressors that are both time persistent and measured with errors. This point is particularly crucial if the right hand side variables are highly time persistent, as is the case here. Under fixed effect estimation, therefore, eliminating heterogeneity bias may come at the cost of exacerbating measurement error bias.[9] To partly deal with this problem we use both pooled 2SLS regressions accounting for some time fixed explanatory variables and we consider a SYS-GMM estimator that, under particular assumptions, controls for unobserved heterogeneity and partly corrects for the deficiencies of the FE estimator.[10]

Endogeneity of other regressors. Although the 2SLS strategy described above addresses the endogeneity of emigration rates, it does not account for the endogeneity of other regressors. For example, the existing literature has studied the impact of human capital and development on institutions, however it is obvious that institutions affect economic performance and the incentives to acquire human capital. In addition, using the lagged dependent in (1) also induces potential biases in the estimation.

To confront the endogeneity issue in a more general way, we will rely on the system-GMM (SYS-GMM) estimator and compare its results with those of the 2SLS method.

[9]For example, this can explain why in the growth literature human capital variables have often been found insignificantly different from zero in panel fixed-effects applications and with negative signs (see Islam, 1995). Hauk and Wacziarg (2009) show that Monte Carlo simulations are in line with these results found in the literature. In addition, even if the model is dynamic they also show that the first-difference GMM estimator does not perform better in terms of bias properties. For example, the Monte Carlo simulations regarding the effect of human capital accumulation on growth display very close results to the fixed effect estimates, suggesting that the weak instrumentation problem may be prevalent in this case.

[10]See also Blundell and Bond (1998), and Bond et al. (2001) that suggest system GMM to be the most appropriate estimator in dynamic panel data model when time series are very persistent.

The SYS-GMM framework accounts for unobservable heterogeneity, endogeneity and persistence of some of the regressors. It allows us to estimate our model with internal instruments only, or with a combination of external and internal instruments. In addition, a sensitivity analysis will also be conducted to check the robustness of the results to the inclusion/exclusion of certain countries (e.g., socialist countries, Sub-Saharan African countries, and oil-exporting countries) whose characteristics may exacerbate reverse causality problems.

2.3 Data

Our data set is a five-year unbalanced panel spanning the period between 1985 and 2010, where the start of the date refers to the dependent variable (i.e., $t = 1985$, $t-1 = 1980$). In our sample, we are considering only developing countries (according to the World Bank Classification), and they enter the panel if they are independent at time $t-1$. The country sample is selected on the basis of the availability of the data described in subsection. Table A3 in Appendix A presents the list of countries in our sample (corresponding to the largest number of observations in panel specifications).

Democracy. Data on democracy are taken from the Freedom House data set, from the POLITY IV data set, and from the Economic Freedom of the World Project.

The Freedom House published the poliitical rights (PR) and civil liberties (CL) indices. They are based on perception measures gathered through expert coding based on news reports, NGOs and think tanks evaluations, and surveys administered to large number of professionals. For the PR index, the questions are grouped into three sub-categories: electoral processes; political pluralism and participation; and functioning of the government. The CL questions are grouped into four subcategories: freedom of expression and belief; association and organization rights; rule of law and personal autonomy; and individual rights. The sum of each country's sub-category scores translates to a rating from 1 to 7, with a higher score indicating more freedom. Following Acemoglu et al. (2008) we transform these indices so that they lie between 0 and 1, with 1 corresponding to the most-democratic set of institutions.

Another measure of democracy from the POLITY IV data set is also considered. Indicators of democracy measure the general openness of political institutions and combines several aspects such as: the presence of institutions and procedures through which citizens can express effective preferences about alternative policies and leaders;

the existence of institutionalized constraints on the exercise of power by the executive power; and the guarantee of civil liberties to all citizens in their daily lives and in acts of political participation. In our data set we consider a composite index (Polity2), that ranges from -10 to + 10. This index is also normalized from 0 to 1. Note that while the "political rights" and "civil liberties" indices are based on public perception measures and can therefore be seen as a reflection of contemporaneous *de facto* institutional quality, the Polity 2 indicator is based on expert coding of legal documents and can therefore be interpreted more as a *de jure* measure.[11]

Finally, we also consider Economic Freedom of the World (EFW), an index which measures the degree to which countries' policies and institutions support economic freedom. Five broad areas are distinguished: (1) size of government; (2) legal structure and security of property rights; (3) access to sound money; (4) freedom to trade internationally; and (5) regulation of credit, labor and business. This index is also normalized between 0 and 1. The ratings are determined by combining real indicators (such as "size of governement", taken from IMF) with answers to survey questions on other modules (such as "independence of the judicial system" taken from perception reports – e.g., the Global Competitiveness Report form the World Economic Forum, or "regulatory restrictions" taken from the World Bank's "Doing Business" database).

Table A1 presents the correlation table between the various institutional indicators. The first three indices (PR, CL, Polity2) exhibit pairwise correlation rates between 0.8 and 0.9; their correlation rate with EFW is around 0.45.

Migration. For emigration data, we use the estimates provided in Bruecker, Capuano, and Marfouk (2013). Focusing on 20 OECD destination countries, they computed emigration stocks and rates of the population aged 25 years and older by gender and educational attainment in 5-year intervals from 1980 to 2010. Data are obtained by harmonizing national censuses and population registers statistics from the receiving countries. On the whole, the 20 destination countries covered represent more than 90 percent of the OECD total immigration stock.

Other data. Data on human capital are based on Barro and Lee (2013). Data on GDP per capita, population, trade, and official development assistance (ODA) are taken from the Penn World Tables and from the World Development Indicators. Data on legal origins are from La Porta et al. (1999), who provide a set of time-invariant

[11] It goes without saying that there is a good deal of discrepancy between de facto and de jure indicators. See Hallward-Driemeier and Pritchett (2011) in the case of the "Doing Business" data.

binary variables characterizing the origin of national law.[12] Ethnic fractionalisation data are taken from Alesina et al. (2003). Latitude and other geograhic and cultural bilateral data from the CEPII database and from Sachs (2003).

Table A2 presents summary statistics for selected variables, calculated considering the largest sample that we use across indicators and estimation techniques.

3 Results

The results are organized in five sub-sections. We first use cross-sectional data to estimate the long-run relationship between emigration and institutional quality depicted in (2) using the OLS and 2SLS regressions with external instruments. Second, we use panel data to estimate the dynamic specification (1) with pooled OLS and 2SLS regressions. Third, we re-estimate the dynamic model using the SYS-GMM technique, combining external and internal instruments. Fourth, we conduct a sensitivity analysis to check the robustness of our results to the exclusion of certain groups of countries (socialist countries, oil-producing countries and sub-Saharan African countries). Finally, we estimate the dynamic model using skill-specific emigration rates to investigate whether the effect of emigration on institutions varies by education level. In the latter two sub-sections, we only rely on the SYS-GMM estimation method. In all cases, the analysis is conducted on four institutional indicators: the Freedom House PR and CL indicators of political rights (PR) and civil liberties (CL), the Polity 2 index, and the index of Economic Freedom of the World (EFW).

3.1 Cross-sectional analysis

Tables 1.a to 1.d report OLS and 2SLS estimates for the long-run specification (2) using data for 2005 for all variables. Standard errors are robust and clustered by country. In OLS regressions (column 1), the estimated coefficient of the emigration rate is positive and statistically significant for each indicator with the exception of the Freedom House Political Rights Index (PR).

In columns 2 to 8, we correct for endogeneity using 2SLS regressions. The emigration rate is instrumented using predicted bilateral stocks generated by a pseudo-

[12] Five systems are distinguished: French, German, British, Scandinavian and Socialist.

gravity model, as explained above.[13] The baseline regression in column 2 shows that the effect of emigration is positive and statistically significant for all indicators. Compared with OLS, the coefficient of the 2SLS regression is larger. We can conclude that the OLS coefficient suffers from a reverse causality bias: emigration rates decrease when institutions improve. Interestingly, the quality of institutions also appears to be positively correlated with our measure of human capital (i.e., the proportion of college graduates in the resident labor force). In columns 3 to 8, we show that our results are robust to the inclusion of additional standard control variables such as absolute geographic variables (latitude, a landlocked dummy variable, area (log) is sq.kms, percentage of malaria area in 1994, percentage of land area in geographical tropics), regional dummies, ethnic fractionalisation and legal origin dummies, and other potential determinants of institutional quality such as GDP per capita, trade (imports + exports as percentage of GDP)[14] and foreign aid (ODA as a percentage of GNI).[15] The inclusion of these control variables does not affect the significance of the emigration coefficient. The coefficient is globally stable, except when we consider the polity2 indicator and control for geographical explanatory variables and regional dummies (colums 4 and 5). We should notice that at least in the case of geographical controls, the quality of the first-stage strongly decreases.[16] As expected, human capital loses significance when GDP is included because of collinearity.

We consider columns 2 and 6 as our preferred specifications. They are suggestive of a positive causal effect of emigration on institutions. Larger effects are found for political institutions than for economic institutions, with long-run effects ranging between 1.4 and 1.6 for the PR index, between 1.2 and 1.3 for the CL index, and between 1.4 and 1.5 for the Polity 2 index. Overall, this means that a 10-percentage point increase in the emigration rate raises standardized democracy indices by 12 to 15 percentage points, that is, by 25 to 30 percent of their standard deviations as reported in Table A1. Regarding the EFW index, the long-run effect ranges only from .3 to .4, implying that a 10 percentage-point increase in emigration raises the index by 3 to .4 percentage points (that is, by 25 to 30% of its standard deviation).

[13] Appendix B describes the model and results are presented in Table A4.

[14] The existing literature has revealed that good institutions are correlated with openness to trade.

[15] Foreign aid can have a negative impact on political institutions as they can lead to rent-seeking activities (Djankov et al., 2008).

[16] The Stock-Yogo weak ID test critical values are respectevey 16.38 and 8.96 for 10% or 15% maximal IV size .

Table 1. Cross-section results
OLS and 2SLS, year 2005

1.a. Dependent = Freedom House Political Rights index (PR)

	(1) OLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS	(7) 2SLS	(8) 2SLS
Human capital	.0147**	.0128**	.0179***	.0224***	.0146**	.0057	.0127**	.0100*
	(.0064)	(.0060)	(.0061)	(.0059)	(.0071)	(.0066)	(.0061)	(.0058)
Total emig. rate	.5670	1.640***	1.395***	1.985***	.994**	1.487***	1.592***	1.568***
	(.3513)	(.5135)	(.4656)	(.7163)	(.5013)	(.5023)	(.5151)	(.4897)
Ethnic fract.			.0800					
			(.1504)					
Log GDP per cap.						.0637*		
						(.0346)		
Trade (% of GDP)							-.0337	
							(.0860)	
Net ODA (% of GNI)								-.2654
								(.1975)
constant	.4025***	.3584***	.1856*	.4798	.2470**	-.0965	.3972***	.3887***
	(.0488)	(.0522)	(.1085)	(.4034)	(.1152)	(.2528)	(.0843)	(.0621)
Legal origin dummies	no	no	yes	no	no	no	no	no
Geographical controls	no	no	no	yes	no	no	no	no
Regional dummies	no	no	no	no	yes	no	no	no
Observations	99	99	97	95	99	97	97	93
KPW F-stat		17.84	20.45	12.25	13.44	16.89	18.22	16.99
Stock-Yogo critical val.								
10% maximal IV size		16.38	16.38	16.38	16.38	16.38	16.38	16.38
15% maximal IV size		8.96	8.96	8.96	8.96	8.96	8.96	8.96

Notes: *** p<0.01; ** p<0.05 and * p<0.1. Robust standard errors clustered by country in parentheses. Col (1) shows OLS results. Col (2) to (8) show 2SLS results; total emig rate is instrumented using geography-based, predicted emigration rates. KPW: Kleibergen-Paap rk Wald F statistics to be compared with the Stock-Yogo critical values for weak instrumentation.

Table 1. Cross-section results (cont'd)
OLS and 2SLS, year 2005

1.b. Dependent = Freedom House Civil Liberties index (CL)

	(1) OLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS	(7) 2SLS	(8) 2SLS
Human capital	.0144***	.0129***	.0162***	.0197***	.0146***	.0063	.0128***	.0105**
	(.0048)	(.0047)	(.0049)	(.0048)	(.0053)	(.0051)	(.0046)	(.0044)
Total emig. rate	.5174***	1.311***	1.107***	1.485***	.9545***	1.158***	1.216***	1.223***
	(.1840)	(.3683)	(.3473)	(.5026)	(.3654)	(.3536)	(.3435)	(.3370)
Ethnic fract.			-.0077					
			(.1065)					
Log GDP per cap.						.0593**		
						(.0271)		
Trade (% of GNP)							-5.8e-04	
							(.0589)	
Net ODA (% of GNI)								-.2557*
								(.1373)
constant	.4456***	.4130***	.3318***	.5416*	.3395***	-.0085	.4259***	.4449***
	(.0371)	(.0397)	(.0808)	(.2790)	(.0841)	(.1950)	(.0648)	(.0462)
Legal origin dummies	no	no	yes	no	no	no	no	no
Geographical controls	no	no	no	yes	no	no	no	no
Regional dummies	no	no	no	no	yes	no	no	no
Observations	99	99	97	95	99	97	97	93
KPW F-stat		17.84	20.45	12.25	13.44	16.89	18.22	16.99
Stock-Yogo critical val.								
10% maximal IV size		16.38	16.38	16.38	16.38	16.38	16.38	16.38
15% maximal IV size		8.96	8.96	8.96	8.96	8.96	8.96	8.96

Notes: *** p<0.01; ** p<0.05 and * p<0.1. Robust standard errors clustered by country in parentheses. Col (1) shows OLS results. Col (2) to (8) show 2SLS results; total emig rate is instrumented using geography-based, predicted emigration rates. KPW: Kleibergen-Paap rk Wald F statistics to be compared with the Stock-Yogo critical values for weak instrumentation.

Table 1. Cross-section results (cont'd)
OLS and 2SLS, year 2005

1.c. Dependent = Polity 2 index

	(1) OLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS	(7) 2SLS	(8) 2SLS
Human capital	.0147***	.0137***	.0174***	.0204***	.0112**	.0122***	.0135***	.0130***
	(.0039)	(.0040)	(.0045)	(.0041)	(.0045)	(.0046)	(.0040)	(.0050)
Total emig. rate	.7920***	1.496**	1.166**	.9870	.7823	1.435**	1.507**	1.389**
	(.2369)	(.6026)	(.5567)	(.7368)	(.5235)	(.6072)	(.6106)	(.5752)
Ethnic fract.			-.1095					
			(.1402)					
Log GDP per cap.						.0112		
						(.0317)		
Trade (% of GNP)							-.0622	
							(.0857)	
Net ODA (% of GNI)								.2173
								(.2719)
constant	.5456***	.5191***	.4589***	.5853	.4985***	.4522**	.5766***	.5101***
	(.0426)	(.0475)	(.1102)	(.3625)	(.1083)	(.2284)	(.0742)	(.0585)
Legal origin dummies	no	no	yes	no	no	no	no	no
Geographical controls	no	no	no	yes	no	no	no	no
Regional dummies	no	no	no	no	yes	no	no	no
Observations	94	94	93	92	94	92	93	88
KPW F stat		16.95	18.99	5.89	20.46	16.36	19.77	17.00
Stock-Yogo critical val.								
10% maximal IV size		16.38	16.38	16.38	16.38	16.38	16.38	16.38
15% maximal IV size		8.96	8.96	8.96	8.96	8.96	8.96	8.96

Notes: *** p<0.01; ** p<0.05 and * p<0.1. Robust standard errors clustered by country in parentheses. Col (1) shows OLS results. Col (2) to (8) show 2SLS results; total emig rate is instrumented using geography-based, predicted emigration rates. KPW: Kleibergen-Paap rk Wald F statistics to be compared with the Stock-Yogo critical values for weak instrumentation.

Table 1. Cross-section results (cont'd)
OLS and 2SLS, year 2005

1.d. Dependent = Economic Freedom of the World index (EFW)

	(1) OLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS	(7) 2SLS	(8) 2SLS
Human capital	.0048**	.0047**	.0060***	.0046***	.0044**	.0012	.0044**	.0025
	(.0020)	(.0019)	(.0020)	(.0017)	(.0021)	(.0020)	(.0019)	(.0022)
Total emig. rate	.3166***	.4112***	.3478***	.3107**	.3823***	.2863***	.3248***	.3670***
	(.1023)	(.1208)	(.1127)	(.1206)	(.1444)	(.0974)	(.1003)	(.0991)
Ethnic fract.			-.0339					
			(.0415)					
Log GDP per c.						.0325**		
						(.0140)		
Trade (% of GNP)							.0184	
							(.0268)	
Net ODA (% of GNI)								-.3347***
								(.1079)
constant	.5908***	.5865***	.5589***	.7840***	.5771***	.3519***	.5813***	.6202***
	(.0169)	(.0168)	(.0454)	(.0738)	(.0417)	(.1073)	(.0223)	(.0184)
Legal origin dummies	no	no	yes	no	no	no	no	no
Geographical controls	no	no	no	yes	no	no	no	no
Regional dummies	no	no	no	no	yes	no	no	no
Observations	75	75	75	73	75	74	74	69
KPW F-stat		16.35	17.29	10.79	12.33	15.43	17.76	14.98
Stock-Yogo critical val.								
10% maximal IV size		16.38	16.38	16.38	16.38	16.38	16.38	16.38
15% maximal IV size		8.96	8.96	8.96	8.96	8.96	8.96	8.96

Notes: *** p<0.01; ** p<0.05 and * p<0.1. Robust standard errors clustered by country in parentheses. Col (1) shows OLS results. Col (2) to (8) show 2SLS results; total emig rate is instrumented using geography-based, predicted emigration rates. KPW: Kleibergen-Paap rk Wald F statistics to be compared with the Stock-Yogo critical values for weak instrumentation.

3.2 Panel analysis with 2SLS

Tables 2.a to 2.d report pooled OLS and 2SLS estimates for the dynamic specification (1). Standard errors are robust and clustered by country. Compared to the cross-section regressions, we now control for the level of the lagged dependent variable and use panel data. The pooled OLS regression in column 1 confirms that institutional indicators are persistent. The coefficient for the lagged dependent usually varies between .7 and .8. This means that it takes 20 to 25 years (4 to 5 periods of 5 years) to reach the long-run level of institutional quality when a shock occurs. The coefficient of human capital remains positive and significant in most regressions, except when we control for regional dummies and GDP per capita. We also identify a positive and significant correlation between institutional quality and emigration, except for the Polity 2 index. As in the previous section, the coefficient increases in 2SLS when emigration is instrumented; this reflects the reverse causality problem of our OLS estimates. Our instrumentation method in pooled OLS builds on Feyrer (2009); it consists in introducing time-dummies and interactions between time-dummies and the log of distance in our pseudo-gravity regression (see Table A4 in Appendix B).

The results in columns 2 to 8 show results that are suggestive of a positive causal effect of emigration on institutions; the effect is robust to the inclusion of control variables. Again, we consider columns 2 and 6 as our preferred specifications. In the panel setting, the short-run effect varies between .4 and .45 for the Freedom House index of political rights, .2 and .25 for the Freedom House index of civil liberties, .3 to .4 for polity 2, and .10 to .12 only for the index of Economic Freedom of the World. Estimation of the dynamic specification confirms that larger effects are found for political institutions than for economic institutions. The long-run effects are obtained by multiplying the short-run coefficient by 4 to 5.[17] Remarkably, they are almost identical to those obtained in the cross-sectional setting; the only difference is that we lose significance for the Polity 2 index. It is worth reminding that Polity 2 captures the quality of *de jure* institutions, whereas the other indicators are mostly based on perceptions and capture the quality of *de facto* institutions.

[17] In the short-run, a 10-percentage point increase in the emigration rate increases the democracy indices by 4 to 6 percent of their standard deviation.

Table 2. Dynamic regressions results
OLS and 2SLS

2.a. Dependent = Freedom House Political Rights index (PR)

	(1) OLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS	(7) 2SLS	(8) 2SLS
PR_{t-1}	.7820***	.7731***	.7673***	.7273***	.7318***	.7494***	.7705***	.7638***
	(.0265)	(.0259)	(.0271)	(.0316)	(.0334)	(.0272)	(.0284)	(.0309)
Human capital$_{t-1}$	.0051**	.0051**	.0058***	.0076***	.0022	.0035	.0052**	.005**
	(.0022)	(.0021)	(.0022)	(.0021)	(.0026)	(.0024)	(.0023)	(.0024)
Total emig. rate$_{t-1}$	.1993**	.4249***	.4038***	.6831***	.2481*	.4097***	.4624***	.4093***
	(.0866)	(.1343)	(.1295)	(.1947)	(.1324)	(.1415)	(.1430)	(.1312)
Ethnic fract.			-.0058					
			(.0397)					
Log GDP pc$_{t-1}$						.0169		
						(.0117)		
Trade (% GDP)$_{t-1}$							-.0022	
							(.0020)	
Net ODA (% GNI$_{t-1}$								.3946
								(1.062)
constant	.0612***	.0545**	.0334	.1142	.0550	-.0575	.0721**	.0598**
	(.0223)	(.0225)	(.0331)	(.1182)	(.0343)	(.0818)	(.0288)	(.0272)
Legal origin dummies	no	no	yes	no	no	no	no	no
Geographical controls	no	no	no	yes	no	no	no	no
Regional dummies	no	no	no	no	yes	no	no	no
Time dummies	yes	yes	yes	yes	yes	yes	yes	yes
Observations	568	568	558	544	568	556	525	488
KPW F-stat		18.06	21.86	15.76	16.49	17.43	17.70	17.72
Stock-Yogo critical val.								
10% maximal IV size		16.38	16.38	16.38	16.38	16.38	16.38	16.38
15% maximal IV size		8.96	8.96	8.96	8.96	8.96	8.96	8.96

Notes: *** p<0.01; ** p<0.05 and * p<0.1. Robust standard errors clustered by country in parentheses. Col (1) shows OLS results. Col (2) to (8) show 2SLS results; total emig rate is instrumented using geography-based, predicted emigration rates. KPW: Kleibergen-Paap rk Wald F statistics to be compared with the Stock-Yogo critical values for weak instrumentation.

Table 2. Dynamic regressions results (cont'd)
OLS and 2SLS

2.b. Dependent = Freedom House Civil Liberties index (CL)

	(1) OLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS	(7) 2SLS	(8) 2SLS
CL_{t-1}	.8052***	.7945***	.7935***	.7474***	.7703***	.7606***	.7991***	.7882***
	(.0252)	(.0254)	(.0248)	(.0307)	(.0281)	(.0280)	(.0285)	(.0303)
Human capital$_{t-1}$	.0048***	.0049***	.0046***	.0062***	.0027	.0027	.0043**	.0034*
	(.0016)	(.0016)	(.0016)	(.0016)	(.0018)	(.0017)	(.0017)	(.0019)
Total emig. rate$_{t-1}$	.1002*	.2499**	.2534**	.3963***	.1720	.2383**	.2162**	.2490***
	(.0572)	(.1005)	(.0996)	(.1483)	(.1055)	(.1001)	(.0977)	(.0928)
Ethnic fract.			-.0188					
			(.0268)					
Log GDP pc$_{t-1}$						.0211**		
						(.0093)		
Trade (% GDP)$_{t-1}$							9.4e-04	
							(.0015)	
Net ODA (% GNI$_{t-1}$								.2611
								(.7365)
constant	.0543***	.0524***	.0604**	.1780**	.0547**	-.0854	.0495**	.0613***
	(.0157)	(.0159)	(.024)	(.0825)	(.0255)	(.0639)	(.0212)	(.0176)
Legal origin dummies	no	no	yes	no	no	no	no	no
Geographical controls	no	no	no	yes	no	no	no	no
Regional dummies	no	no	no	no	yes	no	no	no
Time dummies	yes	yes	yes	yes	yes	yes	yes	yes
Observations	568	568	558	544	568	556	525	488
KPW F-stat		18.29	21.48	16.47	16.54	17.62	18.03	17.88
Stock-Yogo critical val.								
10% maximal IV size		16.38	16.38	16.38	16.38	16.38	16.38	16.38
15% maximal IV size		8.96	8.96	8.96	8.96	8.96	8.96	8.96

Notes: *** p<0.01; ** p<0.05 and * p<0.1. Robust standard errors clustered by country in parentheses. Col (1) shows OLS results. Col (2) to (8) show 2SLS results; total emig rate is instrumented using geography-based, predicted emigration rates. KPW: Kleibergen-Paap rk Wald F statistics to be compared with the Stock-Yogo critical values for weak instrumentation.

Table 2. Dynamic regressions results (cont'd)
OLS and 2SLS

2.c. Dependent = Polity 2 index

	(1) OLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS	(7) 2SLS	(8) 2SLS
Polity2$_{t-1}$	.7688***	.7617***	.7525***	.7342***	.6932***	.7425***	.7664***	.7729***
	(.0318)	(.0304)	(.0334)	(.0371)	(.0420)	(.0327)	(.0312)	(.0311)
Human capital$_{t-1}$	.0060***	.0060***	.0066***	.0079***	.0017	.0050***	.0056***	.0063***
	(.0018)	(.0018)	(.0020)	(.0018)	(.0028)	(.0019)	(.0019)	(.0021)
Total emig. rate$_{t-1}$	.0976	.3037	.2694	.4759	.1553	.2920	.4113*	.1749
	(.1153)	(.2212)	(.2103)	(.3042)	(.2268)	(.2326)	(.2500)	(.1975)
Ethnic fract.			-.0344					
			(.0441)					
Log GDP pc$_{t-1}$						.0111		
						(.0114)		
Trade (% GDP)$_{t-1}$							-.0035	
							(.0025)	
Net ODA (% GNI$_{t-1}$								1.285
								(1.085)
constant	.1072***	.1225***	.1236***	.1658	.1575***	.0540	.1461***	.1124***
	(.0241)	(.0265)	(.0433)	(.1285)	(.0473)	(.0769)	(.0333)	(.0300)
Legal origin dummies	no	no	yes	no	no	no	no	no
Geographical controls	no	no	no	yes	no	no	no	no
Regional dummies	no	no	no	no	yes	no	no	no
Time dummies	yes	yes	yes	yes	yes	yes	yes	yes
Observations	543	543	539	531	543	531	507	470
KPW F-stat		13.76	15.39	5.925	17.05	12.65	15.97	14.70
Stock-Yogo critical val.								
10% maximal IV size		16.38	16.38	16.38	16.38	16.38	16.38	16.38
15% maximal IV size		8.96	8.96	8.96	8.96	8.96	8.96	8.96

Notes: *** p<0.01; ** p<0.05 and * p<0.1. Robust standard errors clustered by country in parentheses. Col (1) shows OLS results. Col (2) to (8) show 2SLS results; total emig rate is instrumented using geography-based, predicted emigration rates. KPW: Kleibergen-Paap rk Wald F statistics to be compared with the Stock-Yogo critical values for weak instrumentation.

Table 2. Dynamic regressions results (cont'd)
OLS and 2SLS

2.d. Dependent = Economic Freedom of the World index (EFW)

	(1) OLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS	(7) 2SLS	(8) 2SLS
EFW_{t-1}	.7491***	.7498***	.7368***	.7232***	.7553***	.7267***	.7385***	.7593***
	(.0290)	(.0279)	(.0264)	(.0290)	(.0325)	(.0283)	(.0280)	(.0304)
Human $capital_{t-1}$	.0020***	.0020***	.0024***	.0020***	.0018***	.0017**	.0020***	.0027***
	(5.0e-04)	(4.9e-04)	(6.5e-04)	(5.0e-04)	(6.9e-04)	(6.6e-04)	(5.3e-04)	(7.6e-04)
Total emig. $rate_{t-1}$	.1331***	.1275***	.1155***	.1070**	.1261**	.1086***	.0864*	.0904**
	(.0315)	(.0435)	(.0443)	(.0485)	(.0566)	(.0392)	(.0464)	(.0382)
Ethnic fract.			-.0148					
			(.0110)					
Log GDP pc_{t-1}						.0027		
						(.0043)		
Trade (% $GDP)_{t-1}$							.0013	
							(8.7e-04)	
Net ODA (% GNI_{t-1}								.9756**
								(.4685)
constant	.1526***	.1525***	.1575***	.2280***	.1468***	.1484***	.1526***	.1403***
	(.0195)	(.0190)	(.0231)	(.0314)	(.0303)	(.0282)	(.0185)	(.0201)
Legal origin dummies	no	no	yes	no	no	no	no	no
Geographical controls	no	no	no	yes	no	no	no	no
Regional dummies	no	no	no	no	yes	no	no	no
Time dummies	yes	yes	yes	yes	yes	yes	yes	yes
Observations	424	424	424	412	424	418	411	382
KPW F-stat		19.80	22.19	20.58	17.85	19.81	24.42	18.38
Stock-Yogo critical val.								
10% maximal IV size		16.38	16.38	16.38	16.38	16.38	16.38	16.38
15% maximal IV size		8.96	8.96	8.96	8.96	8.96	8.96	8.96

Notes: *** p<0.01; ** p<0.05 and * p<0.1. Robust standard errors clustered by country in parentheses. Col (1) shows OLS results. Col (2) to (8) show 2SLS results; total emig rate is instrumented using geography-based, predicted emigration rates. KPW: Kleibergen-Paap rk Wald F statistics to be compared with the Stock-Yogo critical values for weak instrumentation.

3.3 Panel analysis with SYS-GMM

Tables 3.a to 3.d report SYS-GMM estimates for the dynamic specification (1). The advantage of this method is that we can treat all explanatory variables as potentially endogenous. In Column 1, we consider all explanatory variables of interest as predetermined and instrument them using their one-period to third lags. The choice of three lags ensures that the number of instruments is large enough for efficiency, but lower than or equal to the number of country groups.[18] Column 2 considers the same specification but adds external instrument (i.e., we use the emigration rate generated by our pseudo-gravity model as instrument) to the internal ones. Columns 3 to 5 add additional controls as in the previous sub-sections. However we do not control for time invariant variables (fractionalization legal origin dummies, etc.) as in the SYS-GMM, we account for unobserved heterogeneity. Standard errors are robust and clustered by country group.

The coefficient for the lagged dependent is slightly lower than in the 2SLS model; it varies between .7 and .77. This means that it takes 18 to 23 years (3.5 to 4.5 periods of 5 years) to reach the long-run level of institutional quality when a shock occurs. Another difference is that the coefficient of human capital decreases and becomes insignificant in several specifications. Recall that the literature on education and democracy is inconclusive: while Acemoglu et al. (2008) found that education has no explanatory power for democracy, Bobba and Coviello (2007) or Castello-Climent (2008) found a positive and significant effect. Our results suggest that the effect might be there but does not prove to be highly robust to specification choices.

Importantly, our SYS-GMM estimates for emigration remains positive and statistically significant at usual significance levels, even when the Polity 2 index is used. In the SYS-GMM setting, the short-run effect varies between .2 and .25 for the Freedom House index of political rights (PR), .2 and .22 for the Freedom House index of civil liberties (CL), .25 to .35 for polity 2, and .11 to .15 only for the index of Economic Freedom of the World (EFW).

[18] A problem of the GMM estimator is that too many instruments can overfit the endogenous variable. As rule of thumb, the number of instruments should be smaller or equal to the number of countries (see Roodman, 2009). We follow this rule even if sometimes, given few data observations and specifications with additional controls, in the reported regressions the number of instruments is slightly greater than the number of country groups.

Table 3. SYS-GMM dynamic regressions results

3.a. Dependent = Freedom House Political Rights index (PR)

	(1) SYS-GMM	(2) SYS-GMM	(3) SYS-GMM	(4) SYS-GMM	(5) SYS-GMM
PR_{t-1}	.7716***	.7908***	.7697***	.7680***	.7623***
	(.0346)	(.0350)	(.0320)	(.0371)	(.0424)
Human $capital_{t-1}$	.0037*	.0030	.0016	.0040*	.0049
	(.0021)	(.0019)	(.0029)	(.0022)	(.0039)
Total emig. $rate_{t-1}$	.2497**	.2846***	.2592**	.2982**	.2826***
	(.1061)	(.1032)	(.1085)	(.1442)	(.1063)
Log GDP per cap. $_{t-1}$			.0088		
			(.0197)		
Trade (% $GDP)_{t-1}$				-.0053	
				(.0396)	
Net ODA (% GNI_{t-1}					.0478
					(.1813)
constant	.0865***	.0806***	.0245	.0737*	.1046***
	(.0249)	(.0254)	(.1447)	(.0384)	(.0337)
Country fixed effects	yes	yes	yes	yes	yes
Year fixed effects	yes	yes	yes	yes	yes
External Instrument		yes	yes	yes	yes
AR(1) test	1.2e-08	5.8e-09	7.6e-09	1.9e-08	2.6e-07
AR(2) test	.9656	.9714	.9681	.4625	.5385
Hansen J test (p-v)	.1348	.0802	.2320	.2245	.3607
Observations	568	568	556	525	488
N. countries	99	99	97	98	93
N. instr.	57	58	75	75	75

Notes: *** p<0.01; ** p<0.05 and * p<0.1. Robust standard errors clustered by country in parentheses. One-step SYS-GMM estimator. AR(1) and AR(2): p-values of Arellano-Bond test for serial correlations. Hansen J test: p-values for the null hypothesis of instrument validity. Col (1): all variables are instrumented using their own 1st to 3rd lags. Col (2) to (5) adds external instruments (geography-based predicted emigration rate). In addition, SYS-GMM uses 1st differences lagged one period as instruments for the level equations.

Table 3. SYS-GMM dynamic regressions results (cont'd)

3.b. Dependent = Freedom House Civil Liberties index (CL)

	(1) SYS-GMM	(2) SYS-GMM	(3) SYS-GMM	(4) SYS-GMM	(5) SYS-GMM
CL_{t-1}	.7487***	.7784***	.7089***	.7470***	.7464***
	(.0364)	(.0349)	(.0366)	(.0403)	(.0449)
Human capital$_{t-1}$	.0036**	.0025*	-9.7e-04	.0036**	.0040
	(.0016)	(.0014)	(.0024)	(.0018)	(.0027)
Total emig. rate$_{t-1}$	.1918**	.2099**	.2167**	.2243**	.2223***
	(.0822)	(.0852)	(.0906)	(.0999)	(.0858)
Log GDP per cap.$_{t-1}$			.0357**		
			(.0156)		
Trade (% GDP)$_{t-1}$				.0205	
				(.0313)	
Net ODA (% GNI$_{t-1}$					.0561
					(.1121)
constant	.0695***	.0607***	-.1827	.0718**	.0814***
	(.0212)	(.0209)	(.1146)	(.0318)	(.0269)
Country fixed effects	yes	yes	yes	yes	yes
Year fixed effects	yes	yes	yes	yes	yes
External Instrument	no	yes	yes	yes	yes
AR(1) test	2.5e-09	1.3e-09	5.3e-09	3.7e-08	1.4e-07
AR(2) test	.9685	.9551	.9030	.9296	.7428
Hansen J test (p-v)	.0403	.0198	.1381	.1213	.0578
Observations	568	568	556	525	488
N. countries	99	99	97	98	93
N. instr.	57	58	75	75	75

Notes: *** p<0.01; ** p<0.05 and * p<0.1. Robust standard errors clustered by country in parentheses. One-step SYS-GMM estimator. AR(1) and AR(2): p-values of Arellano-Bond test for serial correlations. Hansen J test: p-values for the null hypothesis of instrument validity. Col (1): all variables are instrumented using their own 1st to 3rd lags. Col (2) to (5) adds external instruments (geography-based predicted emigration rate). In addition, SYS-GMM uses 1st differences lagged one period as instruments for the level equations.

Table 3. SYS-GMM dynamic regressions results (cont'd)

3.c. Dependent = Polity 2 index

	(1)	(2)	(3)	(4)	(5)
	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM	SYS-GMM
Polity2$_{t-1}$	.7098***	.7222***	.6981***	.7043***	.7346***
	(.0463)	(.0455)	(.0454)	(.0402)	(.0453)
Human capital$_{t-1}$	.0042**	.0037*	.0025	.0044**	.0052
	(.0021)	(.0021)	(.0027)	(.0022)	(.0038)
Total emig. rate$_{t-1}$	.2515*	.2877**	.2853**	.3551**	.2038*
	(.1294)	(.1411)	(.1434)	(.1503)	(.1235)
Log GDP per cap.$_{t-1}$			.0107		
			(.0175)		
Trade (% GDP)$_{t-1}$				-.0185	
				(.0405)	
Net ODA (% GNI$_{t-1}$					.1373
					(.1656)
constant	.1688***	.1614***	.2588***	.1501***	.1408***
	(.0329)	(.0325)	(.0584)	(.0458)	(.0339)
Country fixed effects	yes	yes	yes	yes	yes
Year fixed effects	yes	yes	yes	yes	yes
External Instrument	no	yes	yes	yes	yes
AR(1) test	2.1e-07	1.9e-07	3.1e-07	9.8e-07	5.4e-06
AR(2) test	.7791	.7789	.7816	.5036	.3877
Hansen J test (p-v)	.1275	.1264	.2539	.3945	.6053
Observations	543	543	531	507	470
N. countries	96	96	94	94	89
N. instr.	57	58	75	75	75

Notes: *** p<0.01; ** p<0.05 and * p<0.1. Robust standard errors clustered by country in parentheses. One-step SYS-GMM estimator. AR(1) and AR(2): p-values of Arellano-Bond test for serial correlations. Hansen J test: p-values for the null hypothesis of instrument validity. Col (1): all variables are instrumented using their own 1st to 3rd lags. Col (2) to (5) adds external instruments (geography-based predicted emigration rate). In addition, SYS-GMM uses 1st differences lagged one period as instruments for the level equations.

Table 3. SYS-GMM dynamic regressions results (cont'd)

3.d. Dependent = Economic Freedom of the World index (EFW)

	(1) SYS-GMM	(2) SYS-GMM	(3) SYS-GMM	(4) SYS-GMM	(5) SYS-GMM
EFW_{t-1}	.6638***	.6632***	.7215***	.6241***	.6719***
	(.0565)	(.0564)	(.0486)	(.0608)	(.0593)
Human $capital_{t-1}$	.0026***	.0026***	.0027**	.0028***	.0041***
	(9.6e-04)	(9.6e-04)	(.0012)	(.0010)	(.0013)
Total emig. $rate_{t-1}$	.1553***	.1527***	.1380***	.1930***	.1157***
	(.0557)	(.0431)	(.0389)	(.0587)	(.0399)
Log GDP per $cap._{t-1}$			-.0042		
			(.0092)		
Trade (% $GDP)_{t-1}$				-.0091	
				(.0159)	
Net ODA (% GNI_{t-1}					.0744
					(.0813)
constant	.2025***	.2030***	.1558***	.1764***	.1446***
	(.0335)	(.0332)	(.0572)	(.0300)	(.0317)
Country fixed effects	yes	yes	yes	yes	yes
Year fixed effects	yes	yes	yes	yes	yes
External Instrument	no	yes	yes	yes	yes
AR(1) test	1.2e-05	1.2e-05	4.1e-06	1.7e-04	2.3e-04
AR(2) test	.0935	.0933	.0839	.1608	.0963
Hansen J test (p-val)	.5533	.5910	.3246	.4800	.6490
Observations	424	424	418	411	382
N. countries	75	75	74	75	69
N. instr.	57	58	75	75	75

Notes: *** p<0.01; ** p<0.05 and * p<0.1. Robust standard errors clustered by country in parentheses. One-step SYS-GMM estimator. AR(1) and AR(2): p-values of Arellano-Bond test for serial correlations. Hansen J test: p-values for the null hypothesis of instrument validity. Col (1): all variables are instrumented using their own 1st to 3rd lags. Col (2) to (5) adds external instruments (geography-based predicted emigration rate). In addition, SYS-GMM uses 1st differences lagged one period as instruments for the level equations.

Estimation of the dynamic specification confirms that larger effects are found for political institutions than for economic institutions. The long-run effects are obtained by multiplying the short-run coefficient by 3.5 to 4.5. Again, they are perfectly in line with those obtained in the cross-sectional and 2SLS settings. The AR(2) test, which tests the null hypothesis that the error term is not second-order serially correlated, and the Hansen J-test of overindentifying restrictions, indicate that the moment conditions are satisfied and the instruments are mostly valid.

3.4 Robustness by sub-sample

The above results suggest that emigration positively affected institutional quality in developing countries between 1985 and 2010. To investigate whether our results could be driven by the inclusion of countries sharing specific characteristics, we re-estimate the model using three alternative samples of countries and relying on the SYS-GMM estimation method with internal instruments or a combination of internal and external instruments. We first exclude socialist countries in columns 1 and 2, defined on the basis of the legal origin dummy. The rationale for doing this is that emigration was legally restricted in these countries prior to the transition while the fall of the Berlin wall drastically affected the evolution of institutions and of emigration patterns. One may also be concerned by the fact that the pre- and post-transition trajectories of human capital have been peculiar in socialist countries (see Acemoglu et al., 2005). Second, we exclude Sub-Saharan African countries in columns 3 and 4. Sub-Saharan African countries are on average less stable politically than the other countries in our sample. Third, we exclude oil-exporting countries in columns 5 and 6. Several studies have pointed out a negative correlation between oil exports (and of natural-resource dependence in general) and democracy (see Ross, 2001, or Tsui 2011).

The results are presented in Table 4 for the Freedom House Political Rights Index (PR). Similar results (available upon request) were obtained for the other indicators. We use a parsimonious specification with the lagged dependent, human capital and the emigration rate. In all cases, the effect of emigration remains positive and significant at the five percent level. The short-run and long-run effects are almost identical to those obtained in the full sample. There is no evidence that the results are driven by heterogeneity between broad groups of countries.

Table 4. Dynamic regressions results by sub-sample

Dependent = Freedom House Political Rights index (PR)

	no socialist countries		no SSA countries		no Oil exporting countries	
	(1)	(2)	(3)	(4)	(5)	(6)
PR_{t-1}	.7482***	.7523***	.7428***	.7630***	.7356***	.7631***
	(.0357)	(.0367)	(.0488)	(.0484)	(.0349)	(.0360)
Human $capital_{t-1}$	.0043	.0040	.0063**	.0059**	.0050**	.0040*
	(.0028)	(.0028)	(.0027)	(.0026)	(.0023)	(.0021)
Total emig. $rate_{t-1}$	.3012**	.3139***	.2825**	.3177***	.2246**	.2717**
	(.1200)	(.1181)	(.1266)	(.1229)	(.1143)	(.1092)
constant	.0877***	.0860***	.1231***	.1156***	.1040***	.0955***
	(.0284)	(.0283)	(.0356)	(.0360)	(.0273)	(.0279)
Country fixed effects	yes	yes	yes	yes	yes	yes
Year fixed effects	yes	yes	yes	yes	yes	yes
External Instrument	no	yes	no	yes	no	yes
AR(1) test	1.7e-07	1.1e-07	1.4e-04	9.4e-05	1.1e-07	5.5e-08
AR(2) test	.6908	.6903	.8875	.8828	.9225	.9139
Hansen J test (p-v)	.2839	.1782	.4137	.3632	.1538	.0937
Observations	469	469	378	378	510	510
N. countries	79	79	67	67	89	89
N. instr.	57	58	57	58	57	58

Notes: *** p<0.01; ** p<0.05 and * p<0.1. Robust standard errors clustered by country in parentheses. One-step SYS-GMM estimator. AR(1) and AR(2): p-values of Arellano-Bond test for serial correlations. Hansen J test: p-values for the null hypothesis of instrument validity. Col (1), (3), (5): all variables are instrumented using their own 1st to 3rd lags. Col (2), (4), (6) adds external instruments (geography-based predicted emigration rate). In addition, SYS-GMM uses 1st differences lagged one period as instruments for the level equations.

3.5 Testing for skill-specific effects

Finally, we investigate whether the effect of emigration is governed by the education level of emigrants (and not only their number). Note that in this section we do not use external instruments (this is to avoid estimating pseudo-gravity models for each education group) and rely instead on the SYS-GMM estimation method with internal instruments only. We start from a parsimonious specification including only

the lagged dependent and the total emigration rate (see Colum 1 of Table 5).[19] We exclude human capital for two reasons: first the significance of this variable was not robust in our benchmark regressions and, second, we want to minimize the risk of collinearity between the shares of college graduates in the resident and emigrant populations. To capture potential heterogeneous effects for different types of migrants (high v. low-skill), we proceed in three steps.

First, we split the total emigration rate by education level. More precisely, we compute skill-specific emigration rates by dividing the number of high-skill (college graduates) and low-skill (with less than college education) emigrants by the total native population (we use the same denominator as in our computation of the total emigration rate). Given that on average, one third of all emigrants have college education, the "high-skill" and "low-skill" emigration rates are respectively equal to one third and to two thirds of the average emigration rate. Note that the "high-skill" and "low-skill" emigration rates are highly correlated with the average emigration rate (the correlation rates vary between 0.86 and 0.98 across samples for the high-skill rate and between 0.87 and 0.99 for the low-skill rate). The results are presented in Colums 2 and 3 of Table 5. The coefficients on the skill-specific emigration rates are significant and higher in magnitude than in Column 1. However, these changes in magnitudes are fully due to scale effects, that is, they are inversely proportional to the scale of the emigration rate included in the regression considered. Indeed, the estimated coefficient is multiplied by 3 when we use the emigration rate of college graduates (one third of the average rate) and by 1.5 when we use the emigration rate of the less educated (two thirds of the average rate).

Second, we interact the total emigration rate with the proportion of highly-educated individuals among emigrants. As can be seen from Column 4 of Table 5, the coefficient on this interaction term, SHM_{t-1}, is not significant.

[19] We only provide results for the Freedom House Political Rights Index (PR). Similar results (available upon request) were obtained for the other indicators.

Table 5. Dynamic regressions results with skill-specific emigration rates

Dependent = Freedom House Political Rights index (PR)

	Full sample: Alternative Specifications				
	(1)	(2)	(3)	(4)	(5)
Total emig. rate$_{t-1}$	.3162***			.3429***	.8027***
	(.0710)			(.0763)	(.2636)
High-skilled emig. rate$_{t-1}$		.8583***			
		(.2105)			
Low-skilled emig. rate$_{t-1}$			.4549***		
			(.1105)		
SHM$_{t-1}$				.0931	
				(.1225)	
Total emig. rate*PR$_{t-1}$					-.6265**
					(.3188)
AR(1) test	6.3e-11	7.2e-11	5.4e-11	1.2e-10	1.9e-10
AR(2) test	.7298	.7172	.7370	.7443	.7359
Hansen J test (p-v)	.1856	.2208	.1365	.4027	.2935
Observations	765	765	765	765	765
N. countries	139	139	139	139	139
N. instr.	40	40	40	57	57

Notes: *** p<0.01; ** p<0.05 and * p<0.1. Robust standard errors clustered by country in parentheses. One-step SYS-GMM estimator. AR(1) and AR(2): p-values of Arellano-Bond test for serial correlations. Hansen J test: p-values for the null hypothesis of instrument validity. All variables are instrumented using their own 1st to 3rd lags. In addition, SYS-GMM uses 1st differences lagged one period as instruments for the level equations. All specifications include the lagged dependent variable (not reported).

Finally, we also consider sub-samples excluding the countries with the highest shares of high/low educated emigrants. Unreported results show that the effect of emigration increases when we drop the quartile of countries with the highest shares of college graduates among emigrants, and conversely if we drop the quartile of countries with the lowest shares. However, we cannot infer from these results that less

educated emigrants are more effective at improving institutions. Indeed, eliminating the top and bottom quartiles based on a different criterion (GDP per capita, or level of democracy) gives similar results: the effet of emigration is greater in poor countries with initially bad institutional quality. Rather, we conclude that the effect of emigration is non linear, a result confirmed in Column 5 of Table 5 where we introduce an interaction term between emigration and the lagged dependent variable. Altogether, the results do not support the existence of heterogeneous effects across skill groups.

4 Conclusion

Emigration affects institutions in developing countries in many ways. By providing people with exit options and a safety-net through remittance income, emigration can lower incentives to voice internally and, eventually, delay democratic reform and political change; on the other hand, emigrants can voice from abroad and support diverse political groups and views at home; they can also contribute to the diffusion of democratic values and norms, be it directly, through return migration and contacts with relatives, or indirectly, through social networks connecting diasporas and home-country populations. Finally, since migration is a non-random process, emigration alters the composition of the home-country population on several dimensions (notably education and ethnicity) that can in turn affect democracy at home.

In this paper we empirically investigate the overall impact of emigration on institutions in a large sample of developing countries. We find that openness to migration (measured by the total emigration rate) contributes to improve institutional quality (as measured by standard indicators of democracy and economic freedom) in the migrants' origin countries. This result is robust to a wide range of specifications and estimation methods as well as to the exclusion of certain groups of countries (e.g., former socialist countries). Remarkably, the cross-sectional estimates are fully in line with the implied long-run relationship obtained from dynamic panel regressions. Also, emigration turns out to be a stronger and more robust determinant of institutions than human capital, at least when using standard measures of human capital such as the Barro and Lee (2013) data. We therefore conclude that emigration to liberal democracies played an important role in determining institutional and political change in developing countries, mostly for the better.

References

[1] Acemoglu, D., S. Johnson and J.A. Robinson (2005): Institutions as a Fundamental Cause of Long-Run Growth and Growth, in Aghion, P. and Durlauff, S., eds.: Handbook of Economic Growth, North Holland. Chapter 6, pp. 385-472.

[2] Acemoglu, D., S. Johnson, J.A. Robinson and P. Yared (2005): From Education to Democracy, American Economic Review, AEA Papers and Proceedings, 95, 2: 44-49.

[3] Acemoglu, D., Johnson S., Robinson J.A., Yared P., (2008) "Income and Democracy", American Economic Review, 98, 3: 808-842.

[4] Alesina, A., Harnoss J., and H. Rapoport (2013): Birthplace Diversity and Economic Prosperity, IZA Discussion Paper No 7568, Institute for the Study of Labor, Bonn.

[5] Alesina, A., A. Devleeschauver, W. Easterly, S. Kurlat and R. Wacziarg (2003): Fractionalization, Journal of Economic Growth, 8: 155-94.

[6] Alesina, A., W. Easterly and J. Matuszeski (2011): Artificial States, Journal of the European Economic Association, 9, 2: 246-77.

[7] Barro, R. and J.W. Lee (2013): A new data set of educational attainment in the world, 1950-2010, Journal of Development Economics, 104: 184-198

[8] Batista, C. and P.C. Vicente (2011): Do migrants improve governance at home? Evidence from a voting experiment, World Bank Economic Review, 25, 1: 77-104.

[9] Beine, M., F. Docquier and H. Rapoport (2001): Brain Drain and Economic Growth: Theory and Evidence, Journal of Development Economics, 64, 1: 275-89.

[10] Beine, M., F. Docquier and H. Rapoport (2007): Measuring international skilled migration: a new database controlling for age of entry, World Bank Economic Review, 21, 2: 249-54.

[11] Beine, M., F. Docquier and H. Rapoport (2008): Brain drain and human capital formation in developing countries: winners and losers, The Economic Journal, 118, 4: 631-52.

[12] Beine, M., F. Docquier and M. Schiff (2013): International migration, transfer of norms and home-country fertility, Canadian Journal of Economics, forthcoming.

[13] Bertoli, S. and F. Marchetta (2013): Bringing it all back home: return migration and fertility choices, World Development, forthcoming.

[14] Blundell, R. and S. Bond (1998): Initial Conditions and Moment Restrictions in Dynamic Panel Data Models, Journal of Econometrics, 87: 115-43.

[15] Bobba , M. and D. Coviello (2007): Weak instruments and weak identification in estimating the effects of education on democracy, Economics Letters, 96: 301-306.

[16] Bond S., Hoeffler A. and J. Temple (2001): GMM Estimation of Empirical Growth Models, CEPR Discussion Paper No. 3048, London: CEPR.

[17] Brucker H., Capuano, S. and A. Marfouk (2013): Education, gender and international migration: insights from a panel-dataset 1980-2010, mimeo.

[18] Castello-Climent A. (2008): On the distribution of education and democracy, Journal of Development Economics, 87: 179-190.

[19] Djankov S., Montalvo J., and M. Reynal-Querol (2008): The curse of aid, Journal of Economic Growth, 13, 3: 169-194.

[20] Djuric, I. (2003): The Croatian Diaspora in North America: Identity, Ethnic Solidarity, and the Formation of a Transnational National Community, International Journal of Politics, Culture and Society, 17, 1: 113-130.

[21] Docquier, F., E. Lodigiani, H. Rapoport and M. Schiff (2011): Emigration and Democracy, World Bank Policy Research Paper No 5557, February.

[22] Docquier, F. and A. Marfouk (2006): International migration by education attainment, 1990-2000; in C. Ozden & M. Schiff (eds): International Migration, Brain Drain and Remittances, New York: McMillan and Palgrave. Chapter 5, pp. 151-99.

[23] Docquier, F. and H. Rapoport (2003): Ethnic discrimination and the migration of skilled labor, Journal of Development Economics, 70, 159-172.

[24] Docquier, F. and H. Rapoport (2012): Globalization, brain drain and development, Journal of Economic Literature, 50, 3: 681-730.

[25] Easterly, W. (2001): The middle class consensus and economic development, Journal of Economic Growth, 6, 4: 317-36.

[26] Eckstein, S. (2009), The Personal is Political: The Cuban Ethnic Electoral Policy Cycle, Latin American Politics and Society, 51, 1: 119-148.

[27] Epstein, G.S., A. Hillman and H.W. Ursprung (1999): The King never emigrates, Review of Development Economics, 3, 2: 107-21.

[28] Fargues, P. (2007): The demographic benefit of international migration: a hypothesis and its application to Middle Eastern and North African Countries; in C. Özden and M. Schiff (eds), International migration, economic development and policy, Washington DC: World Bank and Palgrave Macmillan.

[29] Ferguson, J. (2003): Migration in the Carribean: Haiti, the Dominican Republic and Beyond, Minority Rights Group International.

[30] Feyrer, J. (2009): Trade and Income-Exploiting Time Series in Geography. NBER Working Papers 14910, National Bureau of Economic Research, Inc.

[31] MISSING! Fitzgerald et al. (2013):

[32] Frankel, J. and D. Romer (1999): Does Trade Cause Growth? The American Economic Review, 89, 3: 379-399.

[33] Hallward-Driemeier, M. and L. Pritchett (2011): How business is done and the'doing business'indicators : the investment climate when firms have climate control, Policy Research Working Paper Series 5563, The World Bank.

[34] Haney, P. and W. Vanderbush (1999): The Role of Ethnic Interest Groups in US Foreign Policy: The Case of the Cuban American National Foundation, International Studies Quarterly, 43, 2: 341–361.

[35] Haney, P. and W. Vanderbush (2005): Cuban Embargo: Domestic Politics of American Foreign Policy. Pittsburgh, PA: University of Pittsburgh Press.

[36] Hansen, L.O. (1988): The political and socio-economic context of legal and illegal Mexican Migration to the U.-S. (1942-1984), International Migration, 26, 1: 95-107.

[37] Hauk, W.R. and R. Wacziarg (2009): A Monte Carlo Study of Growth Regressions, Journal of Economic Growth, 14: 103-147.

[38] Head, K., and T. Mayer (2002): Illusory Border Effects: Distance Mis measurement Inflates Estimates of Home Bias in Trade. CEPII Working Paper 2002-01.

[39] Hirschman, A.O. (1970): *Exit, Voice and Loyalty: Responses to Decline in Firms, Organizations and States*, Cambridge, MA: Harvard University Press.

[40] Hirschman, A. O. (1993). Exit, Voice, and the Fate of the German Democratic Republic: An Essay in Conceptual History, World Politics, 45(2), 173-202.

[41] Holland, J. (1999): The American Connection: U.S. Guns, Money and Influence in Northern Ireland, New York: National Book Network.

[42] Islam, N. (1995): Growth Empirics: A panel data approach, Quaterly Journal of Economics, 110, 4:1127-1170.

[43] Katz, E. and H. Rapoport (2005): On human capital formation with exit options, Journal of Population Economics, 18, 2: 267-74.

[44] La Porta, R., Lopez-de-Silanes, F., Shleifer A. and R. Vishny (1999): The quality of government, Journal of Law, Economics and Organization, 15, 1: 222-279.

[45] Mariani, F. (2007): Migration as an antidote to rent-seeking, Journal of Development Economics, 84, 2: 609-30.

[46] Mountford A. (1997): Can a brain drain be good for growth in the source economy?, Journal of Development Economics, 53, 2: 287- 303.

[47] Nickell, S. (1981): Biases in dynamic models with fixed effects, Econometrica, 49, 1417-1426.

[48] Omar Mahmoud, T., H. Rapoport, A. Steinmayr and C. Trebesch (2013): The effect of labor migration on the diffusion of democracy: evidence from a former Soviet Republic, CESifo Working Paper No 4389, September.

[49] Ortega, F. and G. Peri (2013): Migration, Trade and Income, IZA Discussion Paper No. 7325, Institute for the Study of Labor (IZA).

[50] Ragazzi, F. (2009): The invention of the Croatian Diaspora: Unpacking the politics of "diaspora" during the war in Yugoslavia, Center for Global Studies, George Mason University, Working Paper No 10.

[51] Rodrik, D. (2007): One economics, many recipes: globalization, institutions and economic growth, Princeton: Princeton University Press.

[52] Rodrik, D., A. Subramanian and F. Trebbi (2004): Institutions rule: the primacy of institutions over geography and integration in economic development, Journal of Economic Growth, 9, 2: 131-65.

[53] Roodman, D. (2009): A Note on the Theme of Too Many Instruments, Oxford Bulletin of Economics and Statistics, 71, 1: 135-158.

[54] Ross, M.L. (2001): Does oil hinder democracy?, World Politics, 53, 3: 325–61.

[55] Sachs, J. (2003): Institutions don't Rule: Direct Effects of Geography on Per Capita Income. NBER Working Paper No 9490.

[56] Santos Silva, J. M. C., and S. Tenreyro (2011): Poisson: Some convergence issues. Stata Journal, 11(2): 215-225.

[57] Spilimbergo, A. (2009): Foreign students and democracy, American Economic Review, 99, 1: 528-43.

[58] Tsui K.K. (2011): More oil, less democracy: evidence from worldwide crude oil discoveries, The Economic Journal, 121, 551: 89-115.

[59] Vanderbush, W. (2009): Exiles and the Marketing of U.S. Policy toward Cuba and Iraq, Foreign Policy Analysis, 5, 287–306.

[60] Vaubel, R. (2008): A history of thought on institutional competition. In A. Bergh and R. Hoijer, eds.: Institutional Competition, Cheltenham: Edward Elgar.

[61] Wilson, A.J. (1995): Irish America and the Ulster Conflict, 1968-1995, Washington, D.C.: Blackstaff Press.

[62] Wilson, J.D. (2011): Brain-Drain Taxes for Non-benevolent Governments. Journal of Development Economics, 95, 1: 68–76.

Appendix

A. Summary statistics. Table A1 gives the correlation rates between the four institutional indicators used in the paper; correlation rates are computed using pooled data. Table A2 presents summary statistics for selected variables of interest. Summary statistics are computed using observations used in at least one regression. Table A3 lists the countries included in the sample.

Table A1. Correlation rates between institutional indicators

	PR	CL	polity2	EFW
PR	1.000			
CL	.8853	1.000		
polity2	.8444	.7852	1.000	
EFW	.4354	.4984	.4588	1.000

Table A2. Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
PR	568	.454	.326	.0000	1.000
CL	568	.460	.266	.0000	1.000
Polity2	546	.569	.325	.0000	1.000
EFW	434	.576	.115	.1780	.7930
Human capital	568	3.82	4.04	.0610	24.91
Total emigration rate	568	.037	.068	.0001	.4654
Log GDP per capita	556	7.83	.983	5.016	9.743
Trade (as % of GDP)	525	7.29	4.15	.1080	35.87
Net ODA (as % of GNI)	488	.007	.010	-.0001	.0672

Table A3. Countries in the sample

Country	PR/CL	Pol2	EFW	Country	PR/CL	Pol2	EFW	Country	PR/CL	Pol2	EFW	Country	PR/CL	Pol2	EFW
Afghanistan	6	2	0	Dominican R.	6	6	6	Libya	6	6	0	Rwanda	6	6	4
Albania	6	6	4	Ecuador	6	6	6	Lithuania	6	6	3	Senegal	6	6	6
Algeria	6	6	6	Egypt	6	6	6	Malawi	6	6	6	Serbia Mont.	3	3	0
Argentina	6	6	6	El Salvador	6	6	6	Malaysia	6	6	6	Sierra Leone	6	6	6
Armenia	3	3	0	Fiji	6	6	6	Maldives	6	0	0	South Africa	6	6	6
Bangladesh	6	6	6	Gabon	6	6	6	Mali	6	6	6	Sri Lanka	6	6	6
Belize	5	0	5	Gambia	6	6	0	Mauritania	6	6	0	Sudan	6	6	0
Benin	6	6	6	Ghana	6	6	6	Mauritius	6	6	6	Swaziland	6	6	0
Bolivia	6	6	6	Guatemala	6	6	6	Mexico	6	6	6	Syria	6	6	6
Botswana	6	6	6	Guyana	6	6	3	Moldova	3	3	0	Tajikistan	3	3	0
Brazil	6	6	6	Haiti	6	6	6	Mongolia	6	6	0	Tanzania	6	6	6
Bulgaria	6	6	5	Honduras	6	6	6	Morocco	6	6	6	Thailand	6	6	6
Burma	6	6	6	India	6	6	6	Mozambique	6	6	0	Togo	6	6	6
Burundi	6	6	6	Indonesia	6	6	6	Namibia	4	4	4	Tonga	6	0	0
Cambodia	6	4	0	Iran	6	6	6	Nepal	6	6	6	Tunisia	6	6	6
Cameroon	6	6	6	Iraq	6	4	0	Nicaragua	6	6	6	Turkey	6	6	6
Cent. Afr. R.	6	6	5	Jamaica	6	6	6	Niger	6	6	6	Uganda	6	6	6
Chile	6	6	6	Jordania	6	6	6	Pakistan	6	6	6	Ukraine	3	3	3
China	6	6	6	Kazakhstan	3	3	0	Panama	6	6	6	Uruguay	6	6	6
Colombia	6	6	6	Kenya	6	6	6	Papua NG	6	6	5	Venezuela	6	6	6
Congo, DR	6	6	6	Kyrgyzstan	3	3	0	Paraguay	6	6	6	Vietnam	6	6	0
Congo, R	6	6	6	Laos	6	6	0	Peru	6	6	6	Yemen	4	4	0
Costa Rica	6	6	6	Latvia	6	6	3	Philippines	6	6	6	Zambia	6	6	6
C. Ivoire	6	6	6	Lesotho	6	6	0	Romania	6	6	5	Zimbabwe	6	6	6
Cuba	6	6	0	Liberia	6	6	0	Russia	6	6	3	Total	568	543	424

Note: The numbers indicate the observations for each country in the baseline sample in panel regressions

B. Pseudo-gravity model. Table A4 describes the results of the pseudo-gravity model used to predict bilateral migration stocks. Column 1 gives the results of the cross-section regression based on the year 2005. Column 2 gives the results of the panel estimation, in which time-dummies and interactions between time-dummies and distance are included. The panel specification follows Feyrer (2009). Interactions between time dummies and distance account for common shocks in transportation technology (e.g. improvements in aircraft technology have induced more people to move and have reduced long-distance migration costs). As long as changes in transportation technologies are common to all countries, these time series changes will be exogenous with respect to any particular country, but they will have different effects across country pairs, depending on the relative geographic position.

Table A4 shows that geographic characteristics are strong determinants of bilateral migration stocks.[20] As proxies of migration costs, linguistic links favor migration while geographical distances are negatively correlated to bilateral migration stocks. Past guest-worker programs have a positive effect on bilateral migration stocks, as does the population size at origin (bigger countries simply have more migrants in absolute terms).

[20] On the full sample, the simple correlation between the predicted total emigration rate and the observed emigration rate is 0.66, whereas regressing the predicted emigration rate on the actual emigration rate and a constant term yields a (robust) standard error of .073 (coefficient .55) and a R-squared of 0.43.

Table A4. Determinants of bilateral migration stocks (PPML estimation)

	(1)	(2)
Period	2005	1980-2010
$Ling_{ij}$	1.190***	1.050***
	(.1478)	(.1685)
$\ln D_{ij}$	-.7606***	
	(.0740)	
$\ln D_{ij} \times I(1980)$		-1.047***
		(.1788)
$\ln D_{ij} \times I(1985)$		-.9772***
		(.1755)
$\ln D_{ij} \times I(1990)$		-.8073***
		(.0835)
$\ln D_{ij} \times I(1995)$		-.8162***
		(.0799)
$\ln D_{ij} \times I(2000)$		-.8203***
		(.0770)
$\ln D_{ij} \times I(2005)$		-.7919***
		(.0798)
$Guest_{ij}$	1.639***	1.482***
	(.1564)	(.1801)
$\ln P_{i,t}$	.5300***	.5186***
	(0.269)	(.0283)
Constant	7.701***	8.5200***
	(.7644)	(.8342)
Destination dummies	yes	yes
Year dummies	no	yes
N	3,820	22,920

Significant at the level of 10 percent, 5 percent**, 1 percent***. The sample includes 20 destination countries and 191 developed and developing countries of origin. Robust standard errors clustered by country pair in parentheses.