Brain drain and home country institutions*

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

[a] *IRES, Université Catholique de Louvain, IZA, and CReAM*
[b] *CREA, Université du Luxembourg, and Centro Studi Luca d'Agliano*
[c] *Department of Economics, Bar-Ilan University, EQUIPPE, Universités de Lille, and CReAM*
[d] *World Bank, Development Economics Research Group*

Preliminary Draft – August 2009

Abstract

In this paper we empirically investigate the impact of emigration on the quality of institutions in the home country. We consider dynamic-panel regressions and we find a positive effect of both the total emigration rate and the share of tertiary educated workers in the home

*This paper is part of the World Bank Migration and Development Program. We thank Michel Beine, David McKenzie, Caglar Ozden and Robert Vermeulen for useful discussions. We thank also participants at the internal seminar held in Luxembourg University in November 2008 and participants at the second TOM conference, Louvain-la-Neuve, Jan. 2009 for comments on a previous version. We thank Pierre Yared and Cecily Defoort for sharing their data

country, implying that unskilled migration has a positive impact on institutional quality, and skilled migration has an ambiguous impact on institutional quality. Doing some counterfactual simulations, we find that, in general, skilled migration has a positive impact on institutions.

JEL codes: O1, F22.

Keywords: Migration, brain drain, institutions, diaspora effects, voice and exit.

1 Introduction

This paper investigates a neglected aspect of the international migration and development nexus, namely, the relationship between migration and source country institutions and governance (henceforth referred to by the term "institutions"). Many recent studies have emphasized the importance of institutions for economic development and growth (see Acemoglu, 2005, for a detailed survey) and have explored the determinants of institutions, some focusing on geography, and some on policy. For example, Rodrik et al. (2002) show that once institutions are controlled for, i) geography measures have a weak direct effect on income though they have a strong indirect effect through their impact on the quality of institutions, and ii) the direct impact of trade openness is not statistically significant, though its indirect effect - through its impact on institutional quality – is positive as well. Another potentially important determinant of institutions is migration. Indeed, migration offers a safety net and as such can relax economic and political pressures to reform. On the other hand, once abroad, migrants can engage in economic and political activities (e.g., lobbying to encourage or block development aid) that affect the institutional evolution of their home country. In addition, the existence of migration networks abroad increases the home country population's exposure to foreign values and norms as well as the chances of migration for those left behind. Through these exposure and prospective channels, migration can affect the incentive structure faced by individuals at home and modify their choices in terms of education (McKenzie and Rapoport, 2006, Beine et al., 2008), allocation of talent between productive and unproductive activities (Mariani, 2007) or fertility (Fargues, 2006), which can in turn impact on the evolution of home country institutions.

Migration can affect development and growth in source countries in a number of ways. The main channels identified in the literature are remittances (see Rapoport and Docquier, 2006, for a survey), return migration, and the role of migration networks (or diaspora effects). In particular, pre-

vious literature has emphasized the complex relationships between migration and international trade, on the one hand, and FDI, on the other hand, and explored how the relationships between these variables are affected by the skill composition of migration. For example, a series of papers have shown that, contrary to what one would expect from the standard trade-theoretic (Heckscher-Ohlin) framework, trade and migration may be complements rather than substitutes because of migrants' participation in trade networks that reduce transaction and other types of information costs (Gould 1994, Rauch and Trindade, 2002, Rauch and Casella 2003), and because potential migrants may be subject to financing constraints that are alleviated through trade openness (Schiff, 1994). Others have shown that migration and trade will tend to be substitutes (complements) when migration is dominated by skilled (unskilled) labor (e.g., Lopez and Schiff 1998). The same transaction cost argument as outlined above holds for the relationship between migration and FDI (Kugler and Rapoport, 2007; Javorcik et al. 2006; Docquier and Lodigiani, 2009), with a greater impact for skilled than for unskilled migrants. Diasporas also contribute to the creation of scientific networks and favour the diffusion of scientific and technical knowledge (Meyer, 2001, Kerr, 2008, Agrawal et al., 2008), thus fostering technology adoption in the migrants' source countries.

In this paper we focus on the effect of migration on home country institutions. The channels through which migration may impact on the quality of institutions at home are numerous. First and most obviously, emigration and the remittances migrants send back home tend to act as a safety net that can alleviate domestic social, political and economic pressures on source country governments. It may further alleviate these pressures through a reduction in domestic unemployment and an increase in wages. Second, migration is a selective process, with positive self-selection being the rule (Docquier and Marfouk, 2006). Given that the more educated individuals – and the middle class in general – tend to have a higher degree of polit-

ical participation and generally contribute a greater deal to public policy debates, emigration is likely to hurt the quality of domestic institutions and their development as well as the process through which sound policies are being formulated and implemented. Third, in addition to contributing to the public debate, educated individuals also tend to generate a number of more specific externalities, including those related to their impact on other educated individuals' productivity (Lucas, 1988) and to their impact on others' education. Moreover, emigration of entrepreneurs and skilled professionals also tends to result in a substantial loss in governments' net revenues because they tend to pay more taxes than the cost of the services they use. Finally, in addition to their impact through the remittances they send and by taking part in trade and business networks, emigrants also tend to impact their home country through involvement in political networks designed to influence home country politics. This can be achieved directly by expending resources in the home country (e.g., through funding of political parties, the media, various types of associations, etc.) or indirectly through lobbying authorities of host countries in order to affect their policy towards their home country (e.g., channelling of development assistance, or conversely, imposing sanctions). As stated above, the proposed research aims to assess the effect of migration on the source countries' economic and political institutions. This issue is recognized as a crucial one in policy circles (see, e.g., a recent report commissioned by the CIA Strategic Assessment Group (Lahneman, 2005). And yet, with the notable exception of Li and McHale (2006), so far there has been no serious empirical assessment of the channels through which such an impact might operate.

2 An illustrative model

Consider

$$I = \alpha m + \beta h + X \tag{1}$$

where institutional quality I depends on the total emigration rate, m, and on the share of highly skilled human capital in the resident labor force, h. Both m and h depend, on their turn, on $H = \frac{N_s}{N}$, the pre-migration fraction of skilled labor force. In particular:

$$m = \frac{M_s + M_u}{N} = \frac{M_s}{N_s}\frac{N_s}{N} + \frac{M_u}{N_u}\frac{N_u}{N} = m_s H + m_u(1 - H) \tag{2}$$

and

$$h = \frac{N_s - M_s}{N - M_s - M_u} = \frac{\frac{N_s}{N}(1 - m_s)}{\frac{N_s(1-m_s)}{N} + \frac{N_u(1-m_u)}{N}} = \frac{H(1 - m_s)}{H(1 - m_s) + (1 - H)(1 - m_u)} \tag{3}$$

The fraction of skilled labor force before migration may depend on the skilled emigration rate (incentive effect due to migration prospects):

$$H = H(m_s) \tag{4}$$

In order to take into account the impact of skilled and unskilled emigration rates on institutional quality, we consider the derivatives of institutional quality with respect to skilled and unskilled emigration rate. Respectively,

we have:

$$\begin{aligned}
\frac{dI}{dm_s} &= \alpha\frac{dm}{dm_s} + \beta\frac{dh}{dm_s} \\
&= \alpha(H + H'(m_s - m_u)) + \\
&+ \beta\frac{(-H + H'(1-m_s))(H(1-m_s) + (1-H)(1-m_u))}{[H(1-m_s) + (1-H)(1-m_u)]^2} + \\
&+ \beta\frac{(-H(1-m_s))(-H + H'(1-m_s-1+m_u))}{[H(1-m_s) + (1-H)(1-m_u)]^2} \\
&= \alpha H - \beta\frac{H(1-H)(1-m_u)}{[H(1-m_s) + (1-H)(1-m_u)]^2} + \\
&+ H'[\alpha(m_s - mu) + \beta\frac{(1-m_s)(1-m_u) + Hm_s(m_s-m_u)}{[H(1-m_s) + (1-H)(1-m_u)]^2}] \quad (5)
\end{aligned}$$

and

$$\begin{aligned}
\frac{dI}{dm_u} &= \alpha\frac{dm}{dm_u} + \beta\frac{dh}{dm_u} \\
&= \alpha(1-H) + \beta\frac{H(1-H)(1-m_s)}{[H(1-m_s) + (1-H)(1-m_u)]^2} \quad (6)
\end{aligned}$$

Without considering any incentive effects on human capital, e.g. $\frac{\partial H}{\partial m_s} = 0$ (no H'), we can derive some predictions of the model.

Proposition 1: If $\alpha > 0$, $\beta > 0$ and $H < \frac{1}{2}$, then $\frac{dI}{dm_u} > \frac{dI}{dm_s}$, e.g. the marginal impact of the unskilled emigration rate on institutional quality is greater than the marginal impact of the skilled emigration rate.

Proof. Obvious if we consider:

$$\frac{dI}{dm_u} - \frac{dI}{dm_s} = \alpha(1-2H) + \frac{\beta H(1-H)(2-m_s-m_u)}{[H(1-m_s) + (1-H)(1-m_u)]^2} \quad (7)$$

Proposition 2: If $\alpha > 0$, migration without selection improves the quality of institutions, as $\frac{dI}{dm} = \frac{dI}{dm_u}|_m + \frac{dI}{dm_s}|_m = \alpha$

Proposition 3: The marginal impact of skilled emigration rate on institutional quality will be positive, $\frac{dI}{dm_s} > 0$, if $\alpha(1-m) > \beta(1-h)$, where

$(1-m)$ and $(1-h)$ are respectively the average staying rate and the share of low-skilled migrants among residents.

3 Empirical analysis

3.1 General Considerations

The purpose of this section is to empirically investigate the impact of emigration rate on source's countries institutional quality. The first choice for our empirical analysis is the choice between a cross-section or a panel setting. In a cross-sectional dimension, it is possible to use better data both for migration and institutional quality. In particular, for migration, it is possible to use the Docquier Lowell and Marfouk 2007 (DLM) data set, which considers international migration by gender and educational attainment. This data set describes the loss of skilled workers to the OECD for 195 source countries in 1990 and 2000. For institutional quality, the World Bank Governance data by Kaufmann, Kray and Mastruzzi (2006) measures six dimensions of governance from 1996 to 2005: voice and accountability, political stability and absence of violence, government effectiveness, regulatory quality, rule of law, and control of corruption. It covers 213 countries and territories for 1996, 1998, 2000, and annually for 2002-2005. In unreported regressions, we consider OLS regressions using these data sets. We find a significant and positive correlation between the emigration rate and institutional quality indexes, but these regressions suffer from a lot of shortcomings. First, it is difficult to find an appropriate baseline specification, because different economic, political and cultural factors can be important in explaining the quality of institutions. As Alesina et al. (2003) noted, various explanatory variables have been used in the literature on the determinants of institutions, such as log of gdp per capita, legal origin dummies, religious variables, latitude, fractionalisation indices, etc. The main problem with these variables relates to the fact that the pattern of cross-correlations between explanatory variables cannot

be ignored and that in many cases the results of cross-country regressions are sensitive to the econometric specification. For example, as they acknowledge, their index of ethnic fractionalization is highly correlated with latitude and with the log of gdp per capita (which, in addition, is very likely to be endogenous). Moreover, legal origin dummies are highly correlated with religious variables etc. In a panel dimension, instead, it is possible to control for unobservable heterogeneity, and therefore for all time-invariant variables affecting institutional quality. Another problem of cross-sectional analysis refers to endogeneity and reverse causality problems (i.e., bad institutions can cause migration). Attempting to confront the endogeneity issue directly requires finding a suitable instrument. This is not easy in this context. To properly instrument for migration we need a variable that is correlated with the emigration rate but not directly correlated with our endogenous variable, institutional quality. In the migration literature, country's geographical features are often used to instrument for emigration. However, in the institutions literature, the very same geographical characteristics, such as latitude or country size, are also used as determinants of institutions, which would seem to question their theoretical validity as candidate instruments. Therefore it is very difficult to find a proper external instrument for migration. Another problem of the cross-section analysis refers to the fact that institutional quality is a quite persistent variable, therefore a dynamic model would be more suited in order to study the correlation between institution and emigration. Moreover several papers discuss the influence of education on institutional quality, therefore it may be worth to include in our specifications a variable related to education or to human capital. Of course, also this variable would suffer from endogeneity, therefore in a cross-section analysis it would be very difficult to take into account all of these problems. In the next section, we will study the impact of emigration rate on institutional quality using dynamic-panel regressions. In particular, we will use the system-GMM estimator, and we will be able to control for unobservable heterogeneity, to

allow for dynamics, and to account for endogeneity and persistency of some of the variables, using internal instruments.

3.2 Panel analysis

Following Acemoglu et al. (2005), Bobba and Coviello (2007), Amparo Castello-Climent (2008) for studies on democracy and education, and Spilimbergo (2009) for a study on democracy and foreign education, we will study the impact of emigration on institutional quality using dynamic-panel regressions. Using the system-GMM estimator, we will control for unobservable heterogeneity, and we will take into account the endogeneity and persistency of some of the variables. In the following sub-section we will explain the used econometric tecnique.

3.2.1 The econometric model

As in previous studies on democracy and education, and democracy and foreign education, including Acemoglu et al. (2005), Bobba and Coviello (2007), Amparo Castello-Climent (2008), and Spilimbergo (2009), we consider the level of democracy as our dependent variable and we estimate the following dynamic model:

$$\begin{aligned} Democracy_{i,t} &= \beta_0 Democracy_{i,t-1} + \beta_1 h_{i,t-1} + \beta_2 emrate_{i,t-1} + \\ &\quad + \beta_3 X_{i,t-1} + \eta_i + \alpha_t + \varepsilon_{i,t} \end{aligned} \tag{8}$$

where i is the country, t is the period. All explanatory variables are lagged five years. The lagged dependent enters the set of explanatory variables to capture the characteristic of persistency in democracies. The coefficient of interest is β_2 which reflects whether the emigration rate has any positive effects on democracy. The coefficient β_1 captures the effect of human capital on democracy. The coefficient β_3 reflects the importance of other control variables such us the population size (in log) and the structure of the pop-

ulation as in Acemoglu et al. (2005). We also control for time fixed effects, α_t, and country fixed effects, η_t. The advantage of a panel estimation is that it is possible to control for unobservable variables that are country-specific and whose omission, as in the cross-section analysis, can bias the estimated coefficients. Therefore, the results are robust to all country-specific time invariant explanatory variables used in the cross-section literature on institutional quality, including ethnic fractionalisation, religions, legal origins, colonial ties, geographical variables etc.

A general approach to estimate such an equation is to use a transformation that removes unobserved effects and that uses for instrumental variables. Anderson and Hsiao (1992) propose to work with first differences and then to search for instruments. They proposed for the lagged dependent either the two period lagged difference or the two period lagged level of the dependent variable. A generalization of that method was proposed by Arellano-Bond (1991). They suggest using the entire set of instruments in a GMM procedure to reach significant efficiency gains. In the Arellano-Bond method, the first-difference of the explanatory variables are instrumented by the lagged values of the explanatory variables in levels. Under the assumptions that the error term is not serially correlated and that the explanatory variables are weakly exogenous or predetermined (i.e. the explanatory variables are not correlated with future realizations of the error term), the following moment conditions are applied for the first difference equations:

$$E[W_{it-s}.(\Delta\varepsilon_{it})] \quad = \quad 0 \quad for \quad s \geq 2;\ t = 3, ..., T \tag{9}$$

where W_{it-s} are the lagged dependent and all the pre-determined variables in the model. Acemoglu et al.(2005) used this method to study the effect of education on democracy without finding any significant effect. The problem with this method is that, as Bond, Hoeffler and Temple (2001) point out, when time series are persistent, the first-difference GMM estimator can behave poorly: estimates can be seriously biased. To overcome these problems

Bond et al. (2001) suggest to use a more informative set of instruments within the framework developed by Arellano and Bover (1995) and Blundell and Bond (1998). It is widely recognized that democracy varies significantly across countries, but it is quite time-persistent within a country. New results on the relationship between democracy and education were found using Blundell and Bond system GMM estimator.[1] In the case of migration, persistency is also relevant. Therefore, we use the Blundell and Bond system GMM estimator that combines the regression in differences with the regression in levels in a single system. The instruments used in the first differentiated equation are the same as above, but the instruments for the equation in level are the lagged differences of the corresponding variables.

For the level equation the following moment conditions are to be satisfied:

$$E\left[\left(\Delta W_{i,t-1}\right)\left(\eta_i + \varepsilon_{i,t}\right)\right] \;=\; 0 \;\; for \;\; t = 3, 5...T. \tag{10}$$

Condition (10) implies that changes in W are orthogonal to the country fixed effects.

We test the validity of moments conditions by using the test of overidentifying restrictions proposed by Sargan and Hansen and by testing the null hypothesis that the error term is not second order serially correlated. Furthermore, we test the validity of the additional moment conditions associated with the level equation using the Hansen difference test for all GMM instruments.

A particular concern related to this tecnique is the risk of instrument proliferation. In fact, if the use of the entire set of instruments in a GMM context gives significant efficiency gains, on the other hand, a large collection of instruments could overfit endogenous variables as well as weaken the Hansen test of the instruments' joint validity. The instrument proliferation problem is particular important in small samples, but unfortunately there is no formal test to detect it, even if a possible rule of thumb is to keep

[1] Bobba and Coviello (2007), and Amparo Catello (2008).

the number of instruments lower (or equal) the number of groups.[2] In our analysis, we consider the lagged dependent and all the control variables of interest as predetermined, instrumented with "'internal instruments"', using their own lagged one period and further lags, according to the specification, in order to generally have the maximum possible number of instruments, without exceeding the number of groups.

In this setting, thanks to the system GMM estimator, we can handle important modeling concerns such as fixed effects and endogeneity/weakly exogeneity of regressors. Therefore this tecnique is particular suited when it is difficult to find good external instruments, as in our case.

3.2.2 Data

Data are drawn from the following sources:

Democracy

Data on democracy are taken from the Freedom House data set and from the POLITY IV data set.

The Freedom House measures political rights (PR) and civil liberties (CL) using, respectively, an index which ranges from 1 to 7, with an higher score indicating more freedom. The ratings are determined by a list of questions. For the political rights index, for example, the questions are grouped into three sub-categories: electoral process, political plurism and participation, and the functioning of the government. The civil liberties 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. Following Acemoglu et al. (2005) we transform the indexes so that they lie between 0 and 1, with 1 corresponding to the most-democratic set of institutions.

[2]The xtabond2 command, implemented in Stata, gives a warning when instruments exceed the number of groups.

Another measure of democracy from the POLITY IV data set is considered. Indicators of democracy measure the general openness of political institution 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; 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 polity index, that ranges from -10 to + 10. Also this index is normalized from 0 to 1, with 1 corresponding to the most-democratic set of institutions.

Migration and Human Capital

For emigration data, we use the estimates provided in Defoort (2008). Focusing on the six major destination countries (USA, Canada, Australia, Germany, UK and France), she computed skilled emigration stocks and rates by educational attainment from 1975 to 2000 (one observation every 5 years). On the whole, the six destination countries represent about 75 percent of the OECD total immigration stock. However, for some originating countries, the coverage is quite low. For example, Surinamese emigrants mainly live in the Netherlands. About 3 percent of Surinamese emigrants live in the six major receiving countries.

Data on human capital are computed following Defoort (2008), based on Barro and Lee (2001) data set.

Other data

Data on GDP per capita and population data are taken from the PWT and from the World Development Indicators.

Our data set is a five year unbalanced panel spanning the period between 1980 and 2000, where the start of the date refers to the dependent variable (i.e. $t = 1980$, so $t - 1 = 1975$). In our sample, we are considering only developing countries, and they enter the panel if they are independent at

time $t-1$.[3]

3.2.3 Regression results

Tables 1, 2, 3 present our main general results from estimating equation 8 and using the Freedom House PR and CL indicators and the Polity2 measure from the Polity data set. We start considering as variables of interest the lagged dependent, the total emigration rate and the share of tertiary educated workers over the total residence labor force.

Focusing on migration, column 1 of each table shows the pooled OLS relationship between the total emigration rate and democracy by estimating equation 8. The results show a statistically significant and positive correlation between total emigration and democracy, with a coefficient that ranges from 0.314 to 0.237, according to the democracy index considered, and statistically significant at 1-percent level (all the standard errors are robust and clustered by country group). This implies a long-run effect of total emigration on democracy equal to $1.077 \approx 0.320/(1-0.703)$, $0.963 \approx 0.237/(1-0.754)$, $1.082 \approx 0.314/(1-0.710)$, according to the democracy index considered. Column 1 does not control for fixed effect. In column 2, when we control for fixed effect, the coefficient related to the total emigration rate is still positive, but no more significant. 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) showed that the Within estimator is biased of $O(1/T)$). Moreover, both in our fixed effect and pooled OLS estimations, explanatory variables are considered as exogenous. To deal with these problems we use the system GMM estimator that is consistent in dynamic panel estimator, and with "'internal instruments"' we can control for a weak form of exogeneity of all explanatory variables. In fact, we consider the explanatory variables of interest as predetermined, i.e.

[3]The data set used in this paper is an updated version of the data set used by Acemoglu et al. (2005). So many control variables are directly taken from it.

instrumented using their own lagged one period and further lags, in order to use a relevant number of instrument for efficiency reason, but trying to keep the number of instruments lower or at least equal to the number of country groups in all the specification [4]. In column (3) of tables 1, 2, 3, the estimate for the total emigration rate is now positive and again significant at usual statistically level. From column (4-6) of each table, other control variables are introduced in order to check the robustness of our results. In column (4) we add as a regressor the logarithm of population size (lagged), which is positive and statistically significant. Including population size is important to avoid omitted variable bias, because population size appears to be a determinant of institutional quality and it is negatively correlated with the emigration rate (big countries have small emigration rate and vice-versa; therefore including the size of population, we can control that the emigration rate is not just capturing a country-size effect).[5] Following Acemoglu et al. (2005), column (5) includes as additional control variables, the logarithm of population size, the medium age of the population, and also four covariates corresponding to the percentage of the population in different age groups: 0-15, 15-30, 30-45, and 45-60 (the medium age of the population and the four covariates are considered as exogenous to limit the number of instruments). The estimated coefficient of the emigration rate is again positive and statistically significant at 10 and 5 percent, according to the democracy indicator considered. Finally, column (6) controls for gdp per capita (log). The estimated coefficient of the emigration rate is again positive and statistically significant at 10 and 5 percent when considering the civil liberties and polity2 indicators. It looses, instead, its significance when using the political rights indicator. In general, the results on the emigration rate are quite

[4] A problem of the GMM estimator is that too many instruments can over fit the endogenous variable. As rule of thumb, the number of instruments needs to be less or at least equal to the number of groups

[5] That can be the reason why, in some cases, the Hansen test does not pass when we control just for our main variables of interest (specification in column 3)

robust across the different specifications.

Considering column (4) as our best specification, i.e. a parsimonious specification with all statistically significant variables, the coefficients of the total emigration rate are equal to 1.085, 0.786, 1.512 in the short run, and 3.12, 2.18, 3.78 in the long run, when considering, respectively, Political Rights, Civil Liberties and Polity2 as democracy indicators.

If we consider human capital, the coefficient is positive and, in general, statistically significant when using system GMM. Controlling for the structure of the population, human capital is significant at 10 percent if we consider the civil liberties indicator.[6] When we control for gdp per capita, it looses significant, probably because the two variables are highly correlated.[7]

The estimations show also that democracy is very persistent. Considering the first 3 columns, the coefficient on past democracy ranges between the estimated coefficient in pooled OLS, which usually is upward biased, and the estimated coefficient for the fixed effect, which usually displays downward bias. As expected, the unbiased GMM estimator is within this range.

The AR(2) test, the Hansen J test and the difference Hansen test, indicate that the moment conditions are satisfied and the instruments are valid.

From columns 7-9, we consider migration by educational attainment, and starting from our best specification (column 4), we introduce in separate regressions the share of primary, secondary and tertiary educated migrants over total migrants. The coefficient of the share of tertiary educated migrants is negative but not statistically significant. The coefficients of the share of

[6]If we consider the medium age of the population as pre-determined, human capital is in general positive and weakly significant, but in this case we have problem of instrument proliferation, considering instruments till the third lags. Difference-in-Hansen tests of exogeneity of instrument subsets suggest us to treat medium age of the population as exogenous at 5 percent. So it is really "'at the border"'.

[7]In the literature, education is measured as average years of schooling. When we introduce this variable instead of h in specification in column 2, the estimated coefficient is not significant. The estimated coefficient of the emigration rate looses its significance too when considering the PR and CL indicators. It is still significant if considering the Polity2 indicator.

secondary and primary educated migrants are positive, but not statistically significant when using the Freedom House Political Rights and Civil Liberties Indicators. The share of secondary educated migrants is positive and statistically significant, when considering the Polity2 measure, but at the same time the share of tertiary educated residents looses its significance, casting doubts on the fact that the share of secondary educated migrants is likely to capture just an education effect on institutional quality. [8] These results show that the educational structure of migration is not relevant for institutional quality in the originating countries.

3.2.4 Robustness check

In this section we check the robustness of our results considering a weighted emigration rate in the way that we will explain, and a larger sample.

Weighted emigration rate.

Data on migration are taken from Defoort data set (2008). Defoort computed emigration stocks and rates, focusing on the six major destination countries (USA, Canada, Australia, Germany, UK and France). If for some countries, the six destination countries represent a high percent of the OECD total immigration stock; for other countries, the coverage is quite low. Comparing the emigration stock in 2000 in Defoort data set to the one in Docquier and Marfouk data set, which considers 30 OECD destination countries, a variable indicating the percentage of coverage of Defoort data set has been created. We therefore weight the total emigration rate by this coverage measure. Tables 4, 5, 6 show the robustness of our results. Results from the previous section are in fact mainly confirmed.

Larger sample. Previous data on human capital were taken from the widely used Barro and Lee data set. In order to estimate the emigration rate by educational attainment, Defoort (2008) computed also data on labor force by educational attainment for all the countries, considering for developing

[8] Even if the unconditional correlation between h and SMM is 0.3511, so not very high

countries not only data from Barro and Lee, but also from Cohen and Soto and making assumptions where the two data sets were missing. Using this expanded sample, the estimated coefficient of the emigration rate is still positive and highly significant across specifications. The estimated coefficient of human capital is positive, but less significant than before. The different results can be due to a sample effect, now the number of countries is larger than before. Another reason could be linked to data quality and to the use of different data set together.

4 Country specific results

The empirical results show that the total emigration rate has a positive impact on institutional quality in the source country. Therefore, migration without selection improves the quality of institutions (as stated in proposition 2 of our theoretical background), but from our regressions results, we do not know anything about the specific impact of skilled and unskilled emigration. At this purpose, our simple model gives us some predictions in order to understand the marginal impact of emigration by educational attainment. In particular, looking at human capital data, from proposition 1 in section 2, we have that the marginal impact on institutional quality of unskilled emigration is always higher that the marginal impact of skilled emigration, as H is always less that 0.5.

At the same time, recalling proposition 3, the marginal impact of skilled emigration rate on institutional quality is positive if the emigration rate estimated coefficient multiplied by the average staying rate is greater than human capital estimated coefficient multiplied by the share of low-skilled among residents, in particular if $\alpha/\beta > \frac{1-h}{1-m}$. Given our estimated coefficients, it is therefore possible to determine for which country this marginal effect is positive. The benchmark estimated coefficients of emigration rate and human capital are obtained in column (4) of Tables 1, 2, 3, where we

control for population size and human capital data are taken from Barro and Lee data set (and not from constructed data). In these best specifications, we have α equal to 1.085, 0.786, 1.512, and β equal to 0.725, 0.661, 0.573, respectively for political rights, civil liberties and polity2 indicators. Considering these coefficients, we find that the marginal effect of skilled emigration is always positive if we consider the polity2 indicator. For political rights and civil liberties indicators, the effect is positive for most of the countries, except for some small Caribbean countries (see table 10 and figure 1). [9]

In order to evaluate whether skilled emigration has a positive or negative impact on institutional quality, another experiment can be conducted. In particular, the skilled emigration rate can be assumed to be equal to the unskilled emigration rate. This assumption implies a decrease in the skilled emigration rate. Computing the change in institutional quality due to the new simulated value of the skilled emigration rate, it is possible to evaluate if a decrease in the skilled emigration rate improves or decrease institutional quality. In particular, assuming that:

$$I = \alpha m + \beta h + \gamma X \tag{11}$$

and considering $m_s = m_u$, we have:

$$I = \alpha \widetilde{m} + \beta \widetilde{h} + \gamma X \tag{12}$$

where $\widetilde{m} = m_u$ (from equation 2) and $\widetilde{h}$ is the new share of skilled human capital in the residence labor force, due to the decrease in the skilled emigration rate.

[9]For our simulation, we consider the country sample of the benchmark regressions, i.e. 87 countries for PR and CL. We consider the emigration rate and human capital values in year 2000. Our regressions are based on an unbalanced sample, with a minimum of 1 observation per groups and a maximum of 6. Countries enter in our simulation also when they have just 1 observation per period. If the value in 2000 is missing, the constructed value from Defoort (2008) is considered.

The change in institutional quality is given by:

$$\Delta I = \alpha(m_u - m) + \beta(h - \widetilde{h}) \tag{13}$$

From our counterfactual simulations, when considering PR and POL2 indicators ΔI is always negative, implying a negative change in institutional quality when the skilled emigration rate is lower. When considering CL indicator, change in institutional quality is positive for some small Caribbean countries, indicating a better institutional quality with a lower skilled emigration rate (see Tab 11 2).

5 Conclusion

Our analysis shows that the total emigration rate improves institutional quality. Moreover, if unskilled migration has always a positive effect on institutional quality, skilled migration has an ambiguous impact on institutional quality. Using some simulations, we show that skilled emigration rate has a general positive impact on institutional quality (except for some small Caribbean countries).

Table 1: Dependent Variable: Freedom House Political Rights Index

	Pooled OLS (1)	F.E. OLS (2)	SYS GMM (3)	SYS GMM (4)	SYS GMM (5)	SYS GMM (6)	SYS GMM (7)	SYS GMM (8)	SYS GMM (9)
L.prnorm	0.703***	0.386***	0.633***	0.652***	0.602***	0.620***	0.630***	0.645***	0.630***
	(0.0388)	(0.0550)	(0.0645)	(0.0638)	(0.0612)	(0.0612)	(0.0617)	(0.0654)	(0.0607)
L.h	0.830***	-0.429	0.773**	0.725**	0.570	-0.0478	0.993***	0.771*	1.188***
	(0.235)	(0.678)	(0.304)	(0.348)	(0.378)	(0.409)	(0.340)	(0.392)	(0.348)
L.emrtot	0.320***	0.0733	0.317*	1.085***	0.960*	0.208	0.823***	0.659**	0.711**
	(0.0793)	(0.636)	(0.167)	(0.351)	(0.558)	(0.252)	(0.278)	(0.279)	(0.270)
L.lpop				0.0609**	0.0332		0.0433**	0.0290	0.0170
				(0.0263)	(0.0354)		(0.0206)	(0.0231)	(0.0205)
L.medage					0.0790**				
					(0.0314)				
L.agevyoung					3.551				
					(2.379)				
L.ageyoung					1.280				
					(1.885)				
L.agemidage					-0.244				
					(1.311)				
L.ageold					-2.992				
					(2.181)				
L.lrgdpch						0.0760**			
						(0.0311)			
L.SHM							-0.228		
							(0.146)		
L.SMM								0.226	
								(0.187)	
L.SLM									0.183
									(0.116)
Time dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	466	466	466	466	464	416	466	466	466
R-squared	0.607	0.224							
ar1p			0.000	0.000	0.000	0.000	0.000	0.000	0.000
ar2p			0.551	0.565	0.414	0.432	0.648	0.560	0.658
hansenp			0.245	0.392	0.221	0.476	0.390	0.351	0.256
diff. hansen			0.843	0.939	0.641	0.916	0.699	0.894	0.604
Ng		87	87	87	85	81	87	87	87
j			57	74	79	74	76	76	76

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered by country and based on t-statistics in parentheses. One step system GMM estimator. The sample is an unbalanced sample comprising data at five interval between 1980 and 2005. All the variables are treated as pre-determined except for the structure of the population and medage (medium age of the population) that are treated as exogenous. All the pre-determined variables are instrumented for using their own first to third lags from column 3-6. They are instrumented for using their own first to second lags from column 7-9, to avoid instrument proliferation.

Table 2: Dependent Variable: Freedom House Civil Liberties Index

	Pooled OLS	F.E. OLS	SYS GMM	SYS GMM	SYS GMM	SYS GMM	SYS GMM	SYS GMM	SYS GMM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
L.clnorm	0.754***	0.374***	0.601***	0.640***	0.605***	0.590***	0.629***	0.648***	0.632***
	(0.0368)	(0.0546)	(0.0629)	(0.0626)	(0.0593)	(0.0563)	(0.0653)	(0.0641)	(0.0620)
L.h	0.635***	-0.175	0.676***	0.661**	0.470*	0.117	0.768***	0.547**	0.883***
	(0.182)	(0.483)	(0.240)	(0.259)	(0.264)	(0.301)	(0.246)	(0.265)	(0.240)
L.emrtot	0.237***	0.0985	0.315**	0.786***	0.646*	0.338*	0.620***	0.520**	0.591***
	(0.0742)	(0.414)	(0.144)	(0.264)	(0.347)	(0.189)	(0.222)	(0.221)	(0.221)
L.lpop				0.0344**	0.0184		0.0284**	0.0215	0.0135
				(0.0166)	(0.0200)		(0.0139)	(0.0153)	(0.0141)
L.medage					0.0352				
					(0.0223)				
L.age					1.070				
vyoung					(1.726)				
L.age					0.302				
young					(1.358)				
L.age					-0.853				
midage					(0.910)				
L.ageold					-1.489				
					(1.591)				
L.lrgdpch						0.0486**			
						(0.0222)			
L.SHM							-0.175		
							(0.106)		
L.SMM								0.208	
								(0.129)	
L.SLM									0.117
									(0.0865)
Time dumm.	yes	yes	yes	yes	yes	yes	yes	yes	yes
ar1p			0.000	0.000	0.000	0.000	0.000	0.000	0.000
ar2p			0.595	0.557	0.621	0.852	0.520	0.612	0.502
Hansen			0.0660	0.238	0.136	0.231	0.137	0.181	0.142
Diff.			0.362	0.727	0.324	0.419	0.521	0.742	0.61
N.instr.			57	74	79	74	76	76	76
N. count.		87	87	87	85	81	87	87	87
Obs.	466	466	466	466	464	416	466	466	466
R-sq.	0.688	0.322							

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered by country and based on t-statistics in parentheses. One step system GMM estimator. The sample is an unbalanced sample comprising data at five interval between 1980 and 2005. All the variables are treated as pre-determined except for the structure of the population and medage (medium age of the population) that are treated as exogenous. All the pre-determined variables are instrumented for using their own first to third lags from column 3-6. They are instrumented for using their own first to second lags from column 7-9, to avoid instrument proliferation.

Table 3: Dependent Variable: Polity2 Index

	Pooled OLS (1)	F.E. OLS (2)	SYS GMM (3)	SYS GMM (4)	SYS GMM (5)	SYS GMM (6)	SYS GMM (7)	SYS GMM (8)	SYS GMM (9)
L.polity2	0.710***	0.385***	0.584***	0.600***	0.574***	0.593***	0.609***	0.613***	0.585***
	(0.0413)	(0.0560)	(0.0676)	(0.0653)	(0.0682)	(0.0696)	(0.0619)	(0.0679)	(0.0662)
L.h	0.714***	-0.566	0.634**	0.573*	0.491	0.116	0.706**	0.479	0.982***
	(0.214)	(0.666)	(0.307)	(0.341)	(0.425)	(0.397)	(0.317)	(0.332)	(0.364)
L.emrtot	0.314***	0.0244	0.553**	1.512***	1.407**	0.465**	1.029***	0.828**	1.071***
	(0.116)	(0.831)	(0.225)	(0.433)	(0.619)	(0.224)	(0.323)	(0.316)	(0.343)
L.lpop				0.0917***	0.0615		0.0527**	0.0465*	0.0385
				(0.0324)	(0.0417)		(0.0257)	(0.0257)	(0.0245)
L.medage					0.0842**				
					(0.0395)				
L.ageveryyoung					4.705				
					(3.116)				
L.ageyoung					2.158				
					(2.413)				
L.agemidage					0.609				
					(1.625)				
L.ageold					-1.615				
					(2.643)				
L.lrgdpch						0.0566*			
						(0.0288)			
L.SHM							-0.203		
							(0.151)		
L.SMM								0.355**	
								(0.176)	
L.SLM									0.182
									(0.145)
Time dum.	yes	yes	yes	yes	yes	yes	yes	yes	yes
ar1p			0.000	0.000	0.000	0.000	0.000	0.000	0.000
ar2p			0.692	0.687	0.507	0.656	0.735	0.716	0.744
hansenp			0.0394	0.300	0.186	0.321	0.291	0.376	0.265
Diff.			0.07	0.471	0.319	0.903	0.694	0.721	0.797
j			57	74	79	74	76	76	76
N. of count.		82	82	82	82	76	82	82	82
Obs.	450	450	450	450	450	406	450	450	450
R-sq.	0.657	0.424							

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered by country and based on t-statistics in parentheses. One step system GMM estimator. The sample is an unbalanced sample comprising data at five interval between 1980 and 2005. All the variables are treated as pre-determined except for the structure of the population and medage (medium age of the population) that are treated as exogenous. All the pre-determined variables are instrumented for using their own first to third lags from column 3-6. They are instrumented for using their own first to second lags from column 7-9, to avoid instrument proliferation.

Table 4: Dependent Variable: Freedom House Political Rights weighted emrate

	Pooled OLS (1)	F.E. OLS (2)	SYS GMM (3)	SYS GMM (4)	SYS GMM (5)	SYS GMM (6)	SYS GMM (7)	SYS GMM (8)	SYS GMM (9)
L.prnorm	0.702***	0.387***	0.633***	0.649***	0.610***	0.625***	0.625***	0.640***	0.624***
	(0.0385)	(0.0552)	(0.0645)	(0.0632)	(0.0594)	(0.0610)	(0.0611)	(0.0643)	(0.0604)
L.h	0.811***	-0.433	0.762**	0.647*	0.526	-0.0525	0.922***	0.693*	1.100***
	(0.232)	(0.677)	(0.303)	(0.354)	(0.363)	(0.411)	(0.344)	(0.389)	(0.348)
L.emrtotw	0.331***	0.154	0.355*	1.162***	0.977*	0.195	0.915***	0.767**	0.787***
	(0.0823)	(0.608)	(0.184)	(0.363)	(0.540)	(0.287)	(0.299)	(0.292)	(0.284)
L.lpop				0.0641**	0.0335		0.0466**	0.0337	0.0228
				(0.0260)	(0.0339)		(0.0211)	(0.0231)	(0.0208)
L.medage					0.0757**				
					(0.0298)				
L.ageveryyoung					3.375				
					(2.300)				
L.ageyoung					1.141				
					(1.894)				
L.agemidage					-0.236				
					(1.304)				
L.ageold					-3.069				
					(2.157)				
L.lrgdpch						0.0756**			
						(0.0320)			
L.SHM							-0.211		
							(0.150)		
L.SMM								0.234	
								(0.181)	
L.SLM									0.155
									(0.115)
Time dumm.	yes	yes	yes	yes	yes	yes	yes	yes	yes
ar1p			0.000	0.000	0.000	0.000	0.000	0.000	0.000
ar2p			0.548	0.551	0.404	0.433	0.631	0.549	0.635
hansenp			0.266	0.416	0.275	0.478	0.388	0.389	0.258
diff.			0.882	0.953	0.76	0.914	0.687	0.932	0.594
j			57	74	79	74	76	76	76
Ng		87	87	87	85	81	87	87	87
N	466	466	466	466	464	416	466	466	466
R-sq.	0.608	0.224							

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered by country and based on t-statistics in parentheses. One step system GMM estimator. The sample is an unbalanced sample comprising data at five interval between 1980 and 2005. All the variables are treated as pre-determined except for the structure of the population and medage (medium age of the population) that are treated as exogenous. All the pre-determined variables are instrumented for using their own first to third lags from column 3-6. They are instrumented for using their own first to second lags from column 7-9, to avoid instrument proliferation.

Table 5: Dependent Variable: Freedom House CL weighted emrate

	Pooled OLS (1)	F.E. OLS (2)	SYS GMM (3)	SYS GMM (4)	SYS GMM (5)	SYS GMM (6)	SYS GMM (7)	SYS GMM (8)	SYS GMM (9)
L.clnorm	0.752***	0.374***	0.596***	0.631***	0.613***	0.594***	0.619***	0.642***	0.624***
	(0.0367)	(0.0548)	(0.0627)	(0.0616)	(0.0571)	(0.0560)	(0.0644)	(0.0620)	(0.0613)
L.h	0.623***	-0.179	0.689***	0.619**	0.439*	0.101	0.726***	0.489*	0.824***
	(0.179)	(0.481)	(0.240)	(0.260)	(0.253)	(0.296)	(0.248)	(0.258)	(0.240)
L.emrtotw	0.249***	0.173	0.352**	0.864***	0.628*	0.367*	0.711***	0.628***	0.676***
	(0.0759)	(0.392)	(0.152)	(0.280)	(0.332)	(0.207)	(0.241)	(0.235)	(0.236)
L.lpop				0.0374**	0.0170		0.0310**	0.0260*	0.0179
				(0.0171)	(0.0193)		(0.0146)	(0.0156)	(0.0144)
L.medage					0.0320				
					(0.0218)				
L.ageveryyoung					0.898				
					(1.690)				
L.ageyoung					0.205				
					(1.347)				
L.agemidage					-0.848				
					(0.871)				
L.ageold					-1.493				
					(1.544)				
L.lrgdpch						0.0474**			
						(0.0220)			
L.SHM							-0.161		
							(0.109)		
L.SMM								0.215*	
								(0.125)	
L.SLM									0.101
									(0.0871)
Time dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
ar1p			0.000	0.000	0.000	0.000	0.000	0.000	0.000
ar2p			0.603	0.573	0.623	0.847	0.538	0.625	0.523
hansenp			0.0741	0.254	0.138	0.233	0.181	0.197	0.158
difference			0.374	0.772	0.317	0.409	0.683	0.741	0.66
j			57	74	79	74	76	76	76
Number of countries		87	87	87	85	81	87	87	87
N	466	466	466	466	464	416	466	466	466
R-sq.	0.689	0.322							

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered by country and based on t-statistics in parentheses. One step system GMM estimator. The sample is an unbalanced sample comprising data at five interval between 1980 and 2005. All the variables are treated as pre-determined except for the structure of the population and medage (medium age of the population) that are treated as exogenous. All the pre-determined variables are instrumented for using their own first to third lags from column 3-6. They are instrumented for using their own first to second lags from column 7-9, to avoid instrument proliferation.

Table 6: Dependent Variable: Polity2 weighted emrate

	Pooled OLS	F.E. OLS	SYS GMM	SYS GMM	SYS GMM	SYS GMM	SYS GMM	SYS GMM	SYS GMM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
L.polity2	0.709***	0.385***	0.584***	0.597***	0.577***	0.595***	0.602***	0.607***	0.580***
	(0.0413)	(0.0559)	(0.0677)	(0.0651)	(0.0676)	(0.0694)	(0.0625)	(0.0680)	(0.0664)
L.h	0.696***	-0.564	0.623**	0.487	0.453	0.109	0.632*	0.405	0.858**
	(0.212)	(0.667)	(0.302)	(0.340)	(0.411)	(0.394)	(0.320)	(0.322)	(0.356)
L.emrtotw	0.323***	-0.0777	0.562**	1.536***	1.411**	0.473**	1.111***	0.906***	1.126***
	(0.120)	(0.792)	(0.227)	(0.443)	(0.624)	(0.233)	(0.344)	(0.333)	(0.353)
L.lpop				0.0885***	0.0608		0.0539**	0.0475*	0.0416*
				(0.0320)	(0.0419)		(0.0258)	(0.0250)	(0.0245)
L.medage					0.0790**				
					(0.0378)				
L.ageveryyoung					4.378				
					(3.044)				
L.ageyoung					1.935				
					(2.389)				
L.agemidage					0.605				
					(1.610)				
L.ageold					-1.820				
					(2.594)				
L.lrgdpch						0.0554*			
						(0.0293)			
L.SHM							-0.182		
							(0.154)		
L.SMM								0.366**	
								(0.163)	
L.SLM									0.146
									(0.139)
Time dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
ar1p			0.000	0.000	0.000	0.000	0.000	0.000	0.000
ar2p			0.688	0.666	0.494	0.658	0.713	0.700	0.715
hansenp			0.0404	0.327	0.155	0.332	0.303	0.346	0.259
Difference			0.068	0.476	0.212	0.923	0.63	0.635	0.759
j			57	74	79	74	76	76	76
Ng		82	82	82	82	76	82	82	82
Obs.	450	450	450	450	450	406	450	450	450
R-sq.	0.657	0.424							

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered by country and based on t-statistics in parentheses. One step system GMM estimator. The sample is an unbalanced sample comprising data at five interval between 1980 and 2005. All the variables are treated as pre-determined except for the structure of the population and medage (medium age of the population) that are treated as exogenous. All the pre-determined variables are instrumented for using their own first to third lags from column 3-6. They are instrumented for using their own first to second lags from column 7-9, to avoid instrument proliferation.

Table 7: Dependent Variable: Freedom House Political Rights Index (FS)

	Pooled OLS (1)	F.E. OLS (2)	SYS GMM (3)	SYS GMM (4)	SYS GMM (5)	SYS GMM (6)	SYS GMM (7)	SYS GMM (8)	SYS GMM (9)
L.prnorm	0.772***	0.377***	0.671***	0.683***	0.649***	0.626***	0.702***	0.705***	0.710***
	(0.0286)	(0.0480)	(0.0556)	(0.0545)	(0.0504)	(0.0614)	(0.0505)	(0.0506)	(0.0515)
L.hb	0.351**	-0.587	0.490**	0.530**	0.268	-0.256	0.439*	0.445*	0.403*
	(0.176)	(0.551)	(0.221)	(0.239)	(0.284)	(0.357)	(0.232)	(0.239)	(0.227)
L.emrtot	0.256***	0.0568	0.326***	0.406**	0.368	0.343**	0.473***	0.354**	0.409***
	(0.0553)	(0.448)	(0.121)	(0.177)	(0.273)	(0.147)	(0.148)	(0.159)	(0.144)
L.lpop				0.00233	0.00421		0.00863	0.00346	0.00717
				(0.0152)	(0.0186)		(0.0126)	(0.0125)	(0.0115)
L.medage					0.0638***				
					(0.0197)				
L.agevyoung					3.046**				
					(1.533)				
L.ageyoung					2.461*				
					(1.366)				
L.agemidage					-1.067				
					(0.844)				
L.ageold					-0.669				
					(1.418)				
L.lrgdpch						0.0810***			
						(0.0297)			
L.SHM							0.0249		
							(0.0843)		
L.SMM								0.0899	
								(0.138)	
L.SLM									-0.0562
									(0.0760)
Time dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
ar1p			0.000	0.000	0.000	0.000	0.000	0.000	0.000
ar2p			0.704	0.695	0.837	0.684	0.704	0.698	0.715
hansenp			0.116	0.109	0.202	0.299	0.276	0.109	0.269
difference			0.087	0.057	0.655	0.405	0.108	0.058	0.13
j			63	82	87	82	101	101	101
Ng		141	141	141	136	120	141	141	141
N	739	739	739	738	713	585	738	738	738
R-squared	0.658	0.242							

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered by country and based on t-statistics in parentheses. One step system GMM estimator. The sample is an unbalanced sample comprising data at five interval between 1980 and 2005. All the variables are treated as pre-determined except for the structure of the population and medage (medium age of the population) that are treated as exogenous. All the pre-determined variables are instrumented for using their own first to fourth lags.

Table 8: Dependent Variable: Freedom House Civil Liberties Index (FS)

	Pooled OLS (1)	F.E. OLS (2)	SYS GMM (3)	SYS GMM (4)	SYS GMM (5)	SYS GMM (6)	SYS GMM (7)	SYS GMM (8)	SYS GMM (9)
L.clnorm	0.796***	0.387***	0.685***	0.686***	0.636***	0.576***	0.712***	0.714***	0.723***
	(0.0255)	(0.0425)	(0.0486)	(0.0502)	(0.0500)	(0.0520)	(0.0458)	(0.0450)	(0.0455)
L.hb	0.303**	-0.383	0.327*	0.419**	0.148	-0.133	0.348**	0.347*	0.311*
	(0.134)	(0.401)	(0.171)	(0.178)	(0.219)	(0.272)	(0.172)	(0.177)	(0.171)
L.emrtot	0.206***	0.0338	0.318***	0.351**	0.411*	0.362***	0.349***	0.359***	0.375***
	(0.0492)	(0.314)	(0.0921)	(0.142)	(0.220)	(0.118)	(0.121)	(0.130)	(0.113)
L.lpop				-0.00284	0.00238		-0.000162	0.00167	0.00319
				(0.00915)	(0.0117)		(0.00827)	(0.00819)	(0.00770)
L.medage					0.0397**				
					(0.0177)				
L.agevyoung					1.418				
					(1.307)				
L.ageyoung					1.134				
					(1.015)				
L.agemidage					-1.044*				
					(0.593)				
L.ageold					-1.075				
					(0.999)				
L.lrgdpch						0.0616***			
						(0.0224)			
L.SHM							-0.0132		
							(0.0548)		
L.SMM								0.0258	
								(0.0981)	
L.SLM									-0.00365
									(0.0525)
Time dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
ar1p			0.000	0.000	0.000	0.000	0.000	0.000	0.000
ar2p			0.496	0.505	0.524	0.716	0.492	0.500	0.491
hansenp			0.0108	0.0540	0.114	0.136	0.153	0.130	0.0841
Diff.			0.047	0.098	0.614	0.3	0.134	0.31	0.206
j			63	82	87	82	101	101	101
Number of countries		141	141	141	136	120	141	141	141
R-sq.	0.725	0.358							
N	739	739	739	738	713	585	738	738	738

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered by country and based on t-statistics in parentheses. One step system GMM estimator. The sample is an unbalanced sample comprising data at five interval between 1980 and 2005. All the variables are treated as pre-determined except for the structure of the population and medage (medium age of the population) that are treated as exogenous. All the pre-determined variables are instrumented for using their own first to fourth lags.

Table 9: Dependent Variable: Polity2 Index (FS)

	Pooled OLS (1)	F.E. OLS (2)	SYS GMM (3)	SYS GMM (4)	SYS GMM (5)	SYS GMM (6)	SYS GMM (7)	SYS GMM (8)	SYS GMM (9)
L.polity2	0.771***	0.380***	0.626***	0.620***	0.589***	0.606***	0.664***	0.639***	0.665***
	(0.0324)	(0.0502)	(0.0577)	(0.0545)	(0.0560)	(0.0616)	(0.0514)	(0.0529)	(0.0534)
L.hb	0.363**	-0.724	0.557**	0.427*	0.403	-0.0811	0.425*	0.416*	0.426*
	(0.166)	(0.556)	(0.231)	(0.249)	(0.369)	(0.352)	(0.218)	(0.231)	(0.220)
L.emrtot	0.272**	-0.121	0.493**	1.041***	1.212***	0.486**	0.694***	0.696***	0.741***
	(0.109)	(0.799)	(0.204)	(0.313)	(0.393)	(0.219)	(0.216)	(0.263)	(0.243)
L.lpop				0.0627**	0.0625**		0.0308**	0.0299*	0.0315**
				(0.0269)	(0.0302)		(0.0146)	(0.0176)	(0.0146)
L.medage					0.0677**				
					(0.0267)				
L.agevyoung					4.620**				
					(2.167)				
L.ageyoung					2.644				
					(1.782)				
L.agemidage					0.298				
					(1.197)				
L.ageold					1.176				
					(1.800)				
L.lrgdpch						0.0612**			
						(0.0279)			
L.SHM							-0.0556		
							(0.0930)		
L.SMM								0.167	
								(0.135)	
L.SLM									0.0271
									(0.0751)
Time dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
ar1p			0.000	0.000	0.000	0.000	0.000	0.000	0.000
ar2p			0.888	0.930	0.831	0.995	0.962	0.969	0.958
hansenp			0.0317	0.0949	0.199	0.171	0.236	0.288	0.207
Difference			0.009	0.407	0.748	0.33	0.546	0.673	0.541
j			63	82	87	82	101	101	101
Number of countryid		125	125	125	125	109	125	125	125
N	644	644	644	643	643	532	643	643	643
R-squared	0.682	0.433							

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors clustered by country and based on t-statistics in parentheses. One step system GMM estimator. The sample is an unbalanced sample comprising data at five interval between 1980 and 2005. All the variables are treated as pre-determined except for the structure of the population and medage (medium age of the population) that are treated as exogenous. All the pre-determined variables are instrumented for using their own first to fourth lags.

Table 10: Country specific impact of skilled migration on institutional quality ($\frac{\partial I}{\partial m_s} > 0$ if $\alpha/\beta > (1-h)/(1-m)$)

Countries	**(1-h)/(1-m)**	**PR**	**CL**	**POL2**
Afghanistan	0.990	+	+	+
Algeria	0.977	+	+	+
Argentina	0.809	+	+	+
Bangladesh	0.971	+	+	+
Barbados*	1.300	+	-	+
Benin	0.986	+	+	+
Bolivia	0.867	+	+	+
Botswana	0.964	+	+	+
Brazil	0.918	+	+	+
Bulgaria	0.819	+	+	+
Burma (Myanmar)	0.969	+	+	+
Burundi	0.983	+	+	+
Cameroon	0.987	+	+	+
Central African Republic	0.985	+	+	+
Chile	0.854	+	+	+
China	0.975	+	+	+
Colombia	0.920	+	+	+
Congo, Dem. Rep. of the	0.989	+	+	+
Congo, Rep. of the	0.982	+	+	+
Costa Rica	0.838	+	+	+
Croatia	1.013	+	+	+
Czech Republic	0.907	+	+	+
Dominica*	1.429	+	-	+
Dominican Republic	0.973	+	+	+
Ecuador	0.849	+	+	+
Egypt	0.900	+	+	+
El Salvador	1.111	+	+	+
Ethiopia	0.983	+	+	+
Fiji	1.123	+	+	+
Gambia, The	1.011	+	+	+
Ghana	1.006	+	+	+
Guatemala	1.024	+	+	+
Guyana	1.652	-	-	+
Haiti	1.118	+	+	+
Honduras	1.014	+	+	+
Hungary	0.908	+	+	+
India	0.956	+	+	+
Indonesia	0.951	+	+	+
Iran	0.947	+	+	+
Iraq	0.931	+	+	+
Jamaica	1.481	+	-	+
Jordan	0.811	+	+	+

Countries	**(1-h)/(1-m)**	**PR**	**CL**	**POL2**
Kenya	1.004	+	+	+
Lesotho	0.991	+	+	+
Liberia	1.007	+	+	+
Libya	0.899	+	+	+
Malawi	0.998	+	+	+
Malaysia	0.938	+	+	+
Mali	1.003	+	+	+
Mauritania	0.988	+	+	+
Mauritius	1.076	+	+	+
Mexico	1.014	+	+	+
Mozambique	0.999	+	+	+
Nepal	0.978	+	+	+
Nicaragua	0.997	+	+	+
Niger	0.996	+	+	+
Pakistan	0.982	+	+	+
Panama	0.852	+	+	+
Papua New Guinea	0.995	+	+	+
Paraguay	0.923	+	+	+
Peru	0.793	+	+	+
Philippines	0.814	+	+	+
Poland	0.922	+	+	+
Romania	0.927	+	+	+
Rwanda	0.998	+	+	+
Saint Lucia*	1.239	+	-	+
Saint Vincent and the Grenadines*	1.502	-	-	+
Senegal	0.995	+	+	+
Seychelles*	1.193	+	-	+
Sierra Leone	1.010	+	+	+
Slovakia	0.903	+	+	+
South Africa	0.907	+	+	+
Sri Lanka	0.994	+	+	+
Sudan	0.983	+	+	+
Swaziland	0.958	+	+	+
Syria	0.882	+	+	+
Thailand	0.892	+	+	+
Togo	0.990	+	+	+
Trinidad and Tobago	1.275	+	-	+
Tunisia	0.983	+	+	+
Turkey	0.958	+	+	+
Uganda	1.001	+	+	+
Uruguay	0.892	+	+	+
Venezuela	0.826	+	+	+
Vietnam	0.994	+	+	+
Zambia	0.987	+	+	+
Zimbabwe	0.960	+	+	+

Countries	**(1-h)/(1-m)**	**PR**	**CL**	**POL2**
	α/β			
PR	1.497			
CL	1.189			
POL2	2.639			

Table 11: Country specific impact of skilled migration on institutional quality ($m_s = m_u$)

Countries	ΔI **PR**	ΔI **CL**	ΔI **POL2**
Afghanistan	-0.002144	-0.00079435	-0.00543
Algeria	-0.001259	-0.000488105	-0.00312
Argentina	-0.001657	-0.000844303	-0.00346
Bangladesh	-0.000453	-0.000159486	-0.00117
Barbados	-0.022117	-0.002259671	-0.07514
Benin	-0.000789	-0.000281247	-0.00204
Bolivia	-0.00306	-0.001451603	-0.00673
Botswana	-0.000511	-0.000207365	-0.00124
Brazil	-0.000479	-0.000199933	-0.00114
Bulgaria	-0.002433	-0.001206356	-0.00518
Burma (Myanmar)	-0.000528	-0.000214796	-0.00128
Burundi	-0.00056	-0.000214692	-0.0014
Cameroon	-0.001269	-0.000445023	-0.0033
Central African Republic	-0.000477	-0.000173957	-0.00122
Chile	-0.003247	-0.001549008	-0.00711
China	-0.000289	-8.96199E-05	-0.00079
Colombia	-0.0036	-0.001540193	-0.00845
Congo, Dem. Rep. of the	-0.000416	-0.000160491	-0.00103
Congo, Rep. of the	-0.003899	-0.001457525	-0.00984
Costa Rica	-0.004457	-0.002182063	-0.00958
Croatia	-0.004855	-0.001699993	-0.01262
Czech Republic	-0.003118	-0.001378368	-0.00718
Dominica	-0.012484	0.003974561	-0.05932
Dominican Republic	-0.006673	-0.002627727	-0.01641
Ecuador	-0.004547	-0.002209177	-0.00983
Egypt	-0.001846	-0.000815096	-0.00425
El Salvador	-0.004454	-0.001001061	-0.01338
Ethiopia	-0.000727	-0.000261715	-0.00187
Fiji	-0.02095	-0.006145675	-0.05828
Gambia, The	-0.001932	-0.000644535	-0.00513
Ghana	-0.003161	-0.001107998	-0.00821
Guatemala	-0.004169	-0.001408944	-0.011
Guyana	-0.012195	0.014142839	-0.09099
Haiti	-0.012664	-0.003257306	-0.0367
Honduras	-0.004944	-0.001721829	-0.01288
Hungary	-0.005209	-0.002304223	-0.01199
India	-0.000766	-0.000291964	-0.00191
Indonesia	-0.000287	-0.000121971	-0.00068
Iran	-0.003831	-0.001577344	-0.00919
Iraq	-0.002835	-0.001188842	-0.00674
Jamaica	-0.017557	0.005802245	-0.08411
Jordan	-0.00602	-0.00307193	-0.01254

Countries	ΔI **PR**	ΔI **CL**	ΔI **POL2**
Kenya	-0.00257	-0.000912939	-0.00664
Lesotho	-0.00018	-8.15957E-05	-0.00041
Liberia	-0.007173	-0.002598163	-0.01836
Libya	-0.001821	-0.000799345	-0.00421
Malawi	-0.000483	-0.000170274	-0.00125
Malaysia	-0.003236	-0.001334452	-0.00776
Mali	-0.00038	-0.000147649	-0.00094
Mauritania	-0.000804	-0.000300261	-0.00203
Mauritius	-0.008521	-0.002488362	-0.02374
Mexico	-0.001368	-0.000478331	-0.00356
Mozambique	-0.000201	-8.42569E-05	-0.00048
Nepal	-0.000413	-0.000141549	-0.00108
Nicaragua	-0.00991	-0.003758945	-0.02483
Niger	-0.000135	-5.72531E-05	-0.00032
Pakistan	-0.001394	-0.000531522	-0.00348
Panama	-0.013113	-0.006459916	-0.02806
Papua New Guinea	-0.001872	-0.00065676	-0.00486
Paraguay	-0.001174	-0.000512326	-0.00272
Peru	-0.004354	-0.002254979	-0.00896
Philippines	-0.014366	-0.00738719	-0.02974
Poland	-0.005384	-0.002334531	-0.01255
Romania	-0.003341	-0.001428699	-0.00785
Rwanda	-0.000535	-0.00020435	-0.00134
Saint Lucia	-0.016746	-0.00200807	-0.05594
Saint Vincent and the Grenadines	-0.016863	0.006693054	-0.0844
Senegal	-0.001408	-0.000517759	-0.00358
Seychelles	-0.016488	-0.002976246	-0.05186
Sierra Leone	-0.003782	-0.001333723	-0.0098
Slovakia	-0.003407	-0.001520362	-0.0078
South Africa	-0.002972	-0.001301343	-0.00688
Sri Lanka	-0.003691	-0.001362184	-0.00937
Sudan	-0.000515	-0.000193224	-0.0013
Swaziland	-0.0000409	-0.0000055	-0.00013
Syria	-0.002808	-0.001302052	-0.00627
Thailand	-0.000975	-0.000451489	-0.00218
Togo	-0.001616	-0.000575683	-0.00417
Trinidad and Tobago	-0.028675	-0.004876845	-0.09115
Tunisia	-0.001752	-0.000654936	-0.00442
Turkey	-0.000258	-0.00010069	-0.00064
Uganda	-0.001682	-0.000587555	-0.00438
Uruguay	-0.003233	-0.001469075	-0.00732
Venezuela	-0.002402	-0.001198017	-0.00509
Vietnam	-0.00454	-0.001682552	-0.0115
Zambia	-0.001396	-0.000525424	-0.00351
Zimbabwe	-0.002567	-0.001010877	-0.00631

Figure 1: Country specific impact of skilled migration on institutional quality ($\frac{\partial I}{\partial m_s} > 0$ if $\alpha/\beta > (1-h)/(1-m)$)

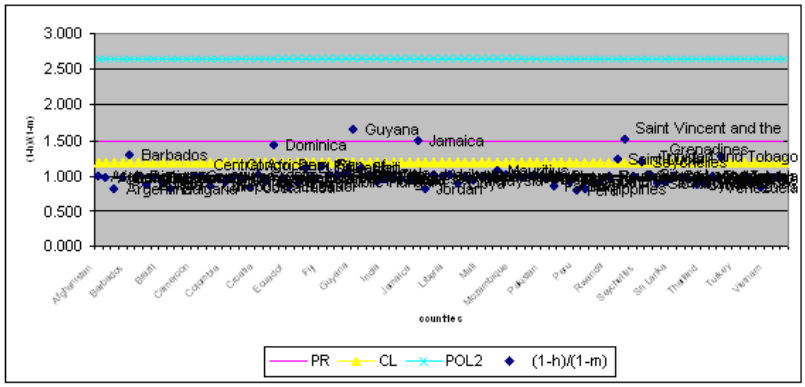
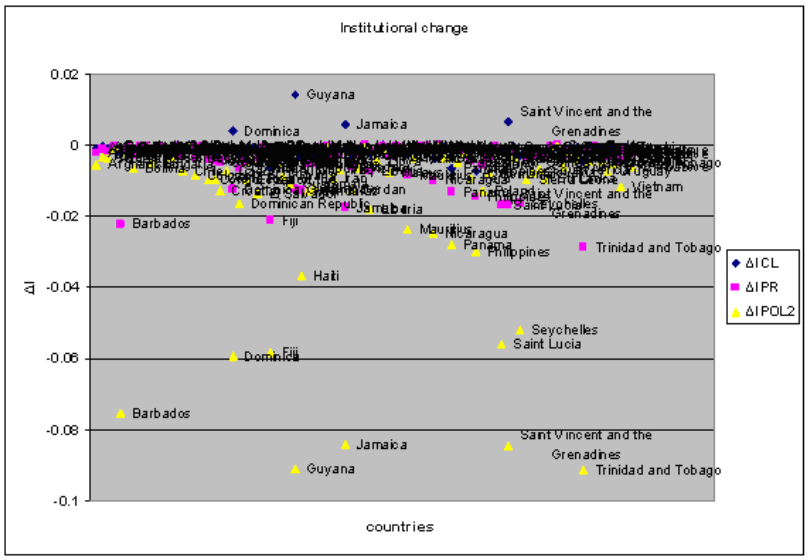


Figure 2: Country specific impact of skilled migration on institutional quality ($m_s = m_u$)



6 References

Acemoglu, D. (2005): Institutions and Growth, in Aghion, P. and Durlauff, S., eds.: Handbook of Economic Growth, North Holland.

Acemoglu, D., S. Johnson, J.A. Robinson and P. Yared (2005), From Education to Democracy, American Economic Review, papers and Prioceedings, 95(2), 44-49.

Agrawal, A.K, Kapur, D. and J. McHale (2008), "Brain drain or brain bank? The impact of skilled emigration on poor country innovation", Mimeo., University of Toronto.

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

Anderson, T.W., and C. Hsiao (1982), "Formulation and Estimation of Dynamic Model Using Panel Data", Journal of Econometrics 18, 67-82.

Arellano, M., and S. R. Bond (1991), "Some Specification Tests for Panel Data: Monte Carlo Evidence and an Application to Employment Equations",

Review of Economic Studies 58, 277-298.

Arellano, M. and O. Bover (1995), "Another Look at the Instrumental Variable Estimation of error-Components Models", Journal of Econometrics 68, 29-51.

Barro, R. and J.W. Lee (2001), International data on educational attainment:updates and implications, Oxford Economic Papers 53, 541-563

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

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

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

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

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

Defoort, C. (2008), Tendances de long terme en migrations internationales. Analyse à partir des 6 principaux pays receveurs, Population-F, 63 (2), 317-352.

Docquier, F., B. L. Lowell and A. Marfouk, (2007), A gendered assesment of the brain drain, IZA Discussion paper 3235.

Docquier, F. and E.Lodigiani (2009): Skilled migration and business networks, Open Economies Review, forthcoming.

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.

Gould, D.M. (1994): Immigrant links to the home country: empirical im-

plications for U.S. bilateral trade flows, Review of Economics and Statistics, 76, 2: 302 16.

Javorcik, B.S., C. Ozden, M. Spatareanu and I.C. Neagu (2006): Migrant Networks and foreign Direct Investment, Mimeo, Rutgers University, November.

Kauffman, D., Kraay, A. & Mastruzzi, M., 2005, 'Governance matters IV: governance indicators for 1996-2004', Mimeo., World Bank

Kerr, W. R. (2008), "Ethnic Scientific Communities and International Technology Diffusion," Review of Economics and Statistics, forthcoming.

Kugler, M., and H. Rapoport (2007), "International labor and capital flows: complements or substitutes?," Economics Letters, 94(2): 155-62.

Lahneman, W.J. (2005). Impact of diaspora communities on global and national politics. Report on survey of the literature, CIA Strategic Assessment Group and University of Maryland, July.

Li, X. and J. McHale (2006). Does brain drain lead to institutional gain? A cross-country empirical investigation. Mimeo. Queen's University.

Lopez, R. and M. Schiff. 1998. 'Migration and the skill composition of the labour force: the impact of trade liberalization in LDCs', Canadian Journal of Economics, 31, 2: 318 36.

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

McKenzie, D. and H. Rapoport (2006): Can migration reduce educational attainments? Evidence from Mexico, World Bank Policy Research Paper No 3952, July.

Meyer, J.B. (2001), "Network approach versus brain drain: Lessons from the diaspora," International Migration, 39, 5: 91-110.

Nickell, S. (1981), "Biases in dynamic models with fixed effects", Econometrica, 49, 1417-1426

Rapoport, H. and F. Docquier (2006): The economics of migrants' remittances, in S.-C. Kolm and J. Mercier Ythier, eds.: Handbook of the

Economics of Giving, Altruism and Reciprocity, North Holland. Chapter 17.

Rauch, J.E. and A. Casella (2003): Overcoming informational barriers to international resource allocation: prices and ties, Economic Journal, 113: 21 42.

Rauch, J.E. and V. Trindade (2002): Ethnic Chinese networks in international trade, Review of Economics and Statistics, 84, 1: 116-30.

Rodrik, D., A. Subramanian and F. Trebbi. 2002. "Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development." CEPR Working Paper No. 3643.

Schiff, M. 1994. "How Trade, Aid and Remittances Affect International Migration," Policy Research Working Paper No. 1376. World Bank.

Spilimbergo, A. (2009): Foreign students and democracy, American Economic Review, 99(1): 52843.

Yasser, A., R. Chami, J. Dagher and P. Montiel (2008): Remittances and institutions: are remittances a curse?, IMF Working Paper No 08-29, February.