A Global Assessment of Human Capital Mobility: the Role of non-OECD Destinations

F. Docquier, C. Özden, Ch. Parsons and E. Artuc

Discussion Paper 2012-22

Institut de Recherches Économiques et Sociales
de l'Université catholique de Louvain

A Global Assessment of Human Capital Mobility: the Role of non-OECD Destinations*

Frédéric Docquier (IRES-UCL, FNRS), Çağlar Özden (World Bank),

Christopher Parsons (University of Oxford), Erhan Artuc (World Bank)

September 2012

Abstract

The discourse concerning the mobility of human capital internationally typically evokes migratory patterns from poorer to relatively more wealthy countries and this focus is strongly reflected in the (brain drain) literature. This emphasis omits an important and as yet understudied aspect of the phenomena however, namely skill transfer to non-OECD and in particular, emerging nations. This paper contributes to the literature by first developing a new dataset of international bilateral migration stocks by gender and education level, which includes both OECD and non-OECD countries as destinations in 1990 and 2000. We then use pseudo-gravity model regressions to impute missing values where data are unavailable, such that we are able to provide, for the first time, a global assessment of human capital mobility. The comprehensiveness of the resulting matrices facilitates a more nuanced definition of emigration rates based on the concept of the natural labour force, which additionally considers both entries and exits of workers.

JEL codes: F22, J61, O15.

Keywords: International migration, labor mobility, brain drain.

*We would like to extend our thanks to the Economic Research Forum for its financial support (contract #2009-057 on "The efficiency and redistributive effects of international labor mobility: a bilateral approach with special focus on MENA migration"). Frederic Docquier also gratefully acknowledges financial support from the Belgian French-speaking Community (convention ARC 09/14-019 on “Geographical Mobility of Factors”). The findings, conclusions and views expressed are entirely those of the authors and should not be attributed to the World Bank, its executive directors or the countries they represent.

1 Introduction

International migration is a powerful force that shapes the distribution of human populations across the globe thereby affecting their social, political, and economic structures. Many aspects of migration have been analyzed by demographers, economists, sociologists, and other social scientists. Migration of the highly skilled (or educated) is one of the most often studied subjects as it has significant growth and development implications for many poor countries that already suffer from low levels of human capital. However, the main focus has always been on the migration of high skilled from poorer to relatively wealthier countries, in tandem with a paucity of available data for migration between developing countries. As a result, existing studies are unable to quantify the extent of the transfer of skills and human capital to non-OECD destinations, especially to less developed countries. These data shortcomings impede many important avenues of research including the determinants of various dimensions of migration patterns (such as gender, age and skill composition) across the whole globe, the reasons behind the emergence and disappearance of important migration corridors, and the analysis of linkages between migration patterns and social and economic development.

A systematic analysis of global migratory patterns by education levels, despite its profound importance and relevance for many research areas and questions, has been conspicuously absent from the literature. More specifically, disentangling the migrant stocks by country of origin, country of destination, gender and education level is crucial to our understanding of the demographic, economic, political and sociological consequences of international labor mobility. The United Nations' International Migrant Stock database provides time series data on the stock of immigrants, by country, but lacks a bilateral dimension. This problem is addressed in the Eurostat database and in Ozden et al. (2011), referred to as OPSW henceforth. The Eurostat database [1] provides data on the size of migration flows, by age, gender and country of citizenship, but solely between EU member states and with missing values. More broadly, OPSW construct five 226x226 comprehensive matrices of origin-destination stocks that correspond to the last five completed census rounds, thereby extending the work of Parsons et al. (2007). However, while OPSW significantly broadens the time, gender and geographical coverage of the available data, different skills or

[1] See http://epp.eurostat.ec.europa.eu/portal/page/portal/population/publications/migration_asylum

education levels are not distinguished.

Another set of studies investigates the education structure of migration. Docquier and Marfouk (2004, 2006) and Dumont and Lemaitre (2004) collect detailed census and register data on immigration from all the host countries of the Organization for Economic Co-operation and Development (referred to as OECD henceforth). Aggregating these numbers allows them to characterize the size and structure of low-skilled and high-skilled emigration stocks to the OECD from all the countries of the world. Docquier, Lowell and Marfouk (2009 - referred to as DLM henceforth) and Dumont, Martin and Spielvogel (2007) introduce the gender breakdown in the above analyzes.

These existing databases of bilateral migrant stocks disaggregated by education level only capture the size and structure of migration to the whole set, or to a large subset of OECD destinations. This is an important limitation, since migration to non-OECD nations is large. Figure 1 shows that the share of non-OECD destination countries in the world immigration stock has gradually decreased since the sixties (from 57 to 49 percent). Nevertheless, non-OECD nations host about half of current international migrants. This share is not homogenous across gender; it is larger for men (51 percent in 2010) than for women (48 percent). Countries such as Russia, Ukraine, India, Pakistan and Ivory Coast attract large numbers of migrants, including many from neighboring countries. As far as high-skilled migration is concerned, countries such as South Africa, the member states of the Gulf Cooperation Council (referred to as GCC henceforth) and some East Asian countries (e.g. Singapore or Hong Kong) are among the most important non-OECD destinations. Omitting these destinations from any analysis, results in an important piece of the global puzzle remains missing, thereby limiting our understanding of the full nature of international human capital mobility.

[INSERT FIGURE 1 AROUND HERE]

In this paper, we develop, for the first time, a global analysis of bilateral migration patterns by gender and for two education levels, i.e. for four labor types. In particular, compared to previous analyses, we account for migration to all non-OECD country destinations by introducing new data and utilizing appropriate estimation methods where data are missing. Our starting point is the work of DLM, which we extend by increasing the number of host countries, adding 46 destinations in 2000 and 31 in 1990 (including 4 OECD new members). Utilizing these observations, we then estimate the size and structure of immigration in the rest of the world using pseudo-gravity regression models.

Comparing migration and labor force data for every country, we are able to refine existing measures of immigration and emigration rates by education levels. In existing studies, emigration rates (and for each labor type) are defined as the number of emigrants (to a limited number of destination countries) divided by the resident labor force at origin. In this paper, instead we characterize the size and structure of the *natural* labor force, i.e. the number of working-age individuals born in a particular origin country, regardless of their current location. Then, for all countries of the world, we are able to express gross emigration stocks and net emigration stocks (emigration minus immigration) as percentages of the natural labor force by labor type.

We find that between 1990 and 2000, migration to non-OECD countries increased at a slower pace (+9.4 percent) than migration to the OECD (+39.2 percent). Nevertheless, these former migrations constitute about 40 percent of the world adult migration stock, and are characterized by both lower shares of women and college graduates (approximately three times less than for migration to OECD countries). This selection on skills is particularly pronounced in the case of least developed countries, increasing with regional income levels and for most global regions between 1990 and 2000. These patterns demonstrate the continued and increasing attractiveness of OECD destinations for high-skilled workers. Conversely however, we find the opposite pattern in terms of the international emigration of females. In other words, although OECD destinations are still broadly favoured by female migrants, the extent of this selection on gender decreased between 1990 and 2000, which highlights the rising appeal of non-OECD destinations to female migrants. Still, emigration to non-OECD countries accounts for about one third of the total brain drain from low-income and the least developed countries. Adding non-OECD destinations increases the high-skilled emigration rate of 32 countries by more than 50 percent. These nations are predominantly those close to South Africa, members of the former Soviet Union or else those that send large numbers of workers to oil producing countries. The influence on our reassessment of female high-skilled emigration based on the additional countries introduced however, is less pronounced given the continued tendency for female migrants to emigrate to OECD nations.

Given the comprehensiveness of our approach, we are also able to construct measures of both net and gross emigration rates. High-income and OECD countries exhibit negative net brain drain rates which show that the incoming pool of educated talent to these regions more than compensates for any skill loss suffered as a consequence of their skilled nationals emigrating abroad. The converse is true of developing regions since although gross and net rates are strongly correlated their net rates are broadly lower. Finally we compare the proportions of educated natives and country residents, the results from which show that globally countries' natural work force is more highly educated than the workforce that resides in that country. In other words, high-skilled immigration to these nations fails to compensate for the skill losses endured when college-educated natives move abroad.

The remainder of this paper is organized as follows. Section 2 describes the methodology and data sources used. Results are then presented in Section 3. Finally, Section 4 concludes.

2 Methodology and definitions

The central goal of this paper is to construct comprehensive bilateral migration matrices by education level and gender for 1990 and 2000, which can be subsequently used to evaluate human capital mobility across the globe and over time. The matrices comprise 195 origin countries (denoted by $j = 1, ..., J$) and include stocks of migrants aged 25 and above. This cutoff is chosen so as to omit students and children since our focus is upon the labor market impact.

The data are disaggregated by gender and two separate education levels. We distinguish males and females, $g = (m, f)$, and two skill types $s = (h, l)$ with $s = h$ for individuals with post-secondary or college education (referred to as the highly skilled), and $s = l$ for less educated individuals (referred to as the low-skilled). In each decade, we thus have migrant stocks of high-skilled males, low-skilled males, high-skilled females and low-skilled females for each bilateral corridor.

Our key variables are denoted as following:

- $M^{jk}_{g,s,t}$: stock of bilateral migrants from country j to country k of gender g and skill s at time t
- $I^{i}_{g,s,t}$: stock of immigrants of type (g, s) to country i at year t
- $E^{i}_{g,s,t}$: stock of emigrants of type (g, s) from country i at year t

- $L^i_{g,s,t}$: (observed) resident labor force of type (g, s) in country i at year t
- $N^i_{g,s,t}$: natural labor force of type (g, s) in country i at year t. This is the number of wrokers from a given country i regardless of their current location.

This section describes the data sources and the methodology used to construct the migration matrices ($M^{jk}_{g,s,t}$ $\forall j, k, g, s, t$) and labor force data.

As explained below, we start with the existing database of DLM, which documents migrant stocks disaggregated by education levels from 195 origin countries to 30 OECD destination countries. We use the same definition as in DLM and add 4 new OECD members and 42 other non-OECD destinations for 2000 (27 non-OECD countries for 1990) for which comparable data could be found. Subsequently, we use the primary data from these 195x76 and 195x61 matrices to predict the size and educational structure of migration stocks for the remaining countries.

For each labor type, aggregation of bilateral migration stocks gives total emigration and immigration of each country:

$$I^i_{g,s,t} \equiv \sum_j M^{ji}_{g,s,t} \tag{1a}$$

$$E^i_{g,s,t} \equiv \sum_k M^{ik}_{g,s,t}. \tag{1b}$$

We will then use data on the educational and gender structure of the labor force to identify the vectors of $L^i_{g,s,t}$ and $N^i_{g,s,t}$ for all i, g, s, t. By definition, the observed resident labor force of type (g, s) in country i, $L^i_{g,s,t}$, is equal to the *non-migrant* labor force (people who have never moved) plus immigrants. Similarly, natural labor force of type (g, s) in country i , $N^i_{g,s,t}$, is equal to *non-migrant* labor force plus the emigrants. Thus, we can state that the non-migrant labor can be expressed as either of the following expressions (residents minus immigrants or naturals minus emigrants):

$$L^i_{g,s,t} - I^i_{g,s,t} = N^i_{g,s,t} - E^i_{g,s,t} \tag{2}$$

Before delving into the details of the empirical exercise and the analysis of our data, we first present summary statistics of the numbers of high-skilled migrants in the database in Table 1. We distinguish between migration to OECD and non-OECD countries and between raw data and estimated/imputed data. For each year, the migrant stock in the 34 OECD countries is shown in the second column. There are

59.3 million migrants above age 25 in 2000, of which 20.9 million (35 percent) have college education, and 30.3 million (51 percent) are women. For 1990, we identify 42.6 million migrants including 30 percent highly educated and 51 percent women.

The third and fourth columns show the data obtained or estimated for non-OECD countries. There are 41.2 million migrants, of which 5.3 million (13 percent) are highly educated and 18.8 million (46 percent) are female. For 1990, we identify 37.7 million migrants, including 10 percent highly educated and 44 percent women. In comparison with OECD destinations, the shares of both the high-skilled and female migrants in non-OECD countries are lower.

Finally, for completeness, the fifth and sixth rows present the numbers and the proportions of migrants predicted for the unobserved corridors. For 2000, imputed values represent 22.9 million migrants in the 119 destination countries for which actual data are not available. Although imputed values account for 22.8 percent of the total migration stock in 2000, the share is around 11 percent for college-educated migrants (around 15 percent in 1990). In other words, since between 85-89% of college-educated migrants are captured by our raw data, we believe that our imputation strategy should not adversely effect our overall measurement of high-skilled emigration to the extent that our conclusions are excessively diluted.

Overall, the migration matrices identify 100.5 million migrants (age 25+) in 2000 which represents about 56.7 percent of the 177.4 million migrants (age 0+) recorded in the United Nations database and 62.8 percent of the 160.1 million migrants (again age 0+) recorded in OPSW for those 195 countries that appear in our matrices;[2] 26.2 million of this migrant stock have college education and 49.0 million are women. For 1990, we identify 80.2 million migrants (aged 25+), including 16.4 million high-skilled migrants and 38 million women. Our data show that the overall migrant stock increased by 25.3 percent between 1990 and 2000, while the stock of high-skilled migrants increased by 60.1 percent. As a result, the share of high-skilled in the overall migrant stock increased from 20.4 percent to 26.1 percent. The share of women increased from 47.3 percent to 48.8 percent, a result mainly driven by the increased feminization of migration to non-OECD countries.

[INSERT TABLE 1 AROUND HERE]

[2] There are differences between OPSW and the United Nations database. For example, OPSW remove refugees wherever possible from their data since their primary focus is upon economic migration.

2.1 Migration data for OECD countries

Our starting point in the construction of our matrices is the DLM (2009) dataset, that comprises a collection of census and register data by country of birth, education level and gender for OECD countries in 1990 and 2000. The original DLM dataset omitted data for member states that subsequently joined the OECD in 2010 however (namely Chile, Estonia, Israel and Slovenia), so in this paper we augment the original OECD data from the DLM dataset with census data pertaining to these newer members. As a result, our set of OECD countries includes all 34 current members for both 1990 and 2000.

DLM enumerates stocks of migrants living in a destination country at the time of census as opposed to flows that are observed between two points in time. For reasons of consistency and comparability, the methodological choices made in DLM guide our current work. The four main principles are the following:

- 195 origin countries are distinguished: starting with the 192 UN member states; we aggregate South Korea and the Democratic People's Republic of Korea since some destination countries only provide the total number of Koreans; Serbia and Montenegro are treated as a single entity and the Holy See, Taiwan, Hong Kong, Macao and the Palestinian Territories are added as individual entries to the country list.

- Migration is measured on the basis of country of birth as opposed to citizenship, since our goal is to have a consistent defintion over time. Whereas individuals' country of birth is predominanty time invariant and independent of the variation in laws regarding citizenship within and across countries, the concept of citizenship conversely changes with naturalizations. Furthermore, many destination countries grant citizenship selectively to migrants from certain countries, significantly biasing the overall migration data based on citizenship status.

- Only adult migrants aged 25 and above are recorded. This measure therefore excludes both students, who temporarily relocate to complete their education, as well as children who accompany their parents abroad. This is a superior measure when wishing to examine the economic and labour market effects of migration.

- Along with the gender dimension, two separate levels of education are distinguished. High-skill migrants include those with college or post-secondary education. Low-skill migrants include all of those with a level of schooling up to and including an upper-secondary education.[3]

As shown in Table 1, the OECD data allow us to characterize the education level, origin and destination of about 59.3 million migrants in 2000 and 42.6 million migrants in 1990. About 16.9 million of the 20.9 million high-skill migrants in the OECD countries are concentrated in only 5 destination countries: the U.S. (10.3 million), Canada (2.7 million), Australia (1.6 million), the United Kingdom (1.2 million) and Germany (1.2 million).

2.2 Migration data for non-OECD countries

We further supplement our expanded data collection for our OECD destinations with 42 non-OECD countries in 2000 and 27 countries in 1990 adhering to the same methodological principles and definitions as in DLM. This non-OECD group includes the following countries, where a superscript * indicates that data are only available for 2000 and are missing for 1990:

- 8 European Union member states: Bulgaria, Croatia, Cyprus, Latvia, Lithuania, Macedonia, Malta, and Romania,
- 13 Central and South American countries: Argentina, Belize*, Bolivia*, Brazil, Colombia, Costa Rica, the Dominican Republic*, Honduras*, Nicaragua*, Panama*, Paraguay*, Trinidad and Tobago* and Venezuela,
- 14 Asian countries: Bahrain, Belarus, Hong Kong*, Iraq*, Kuwait, Kyrgyzstan*, Malaysia*, Mongolia*, Oman, the Philippines, Qatar, Saudi Arabia, Singapore and the United Arab Emirates,
- 7 African countries: Guinea*, the Ivory Coast, Kenya, Morocco*, Rwanda, South Africa and Uganda.

[3]Note that DLM disaggregated low-skill migrants in two categories, those with upper-secondary education and those with less (including low-secondary, primary or no schooling). In this paper, we aggregate these two categories for estimation purposes.

The data sources for these destination countries, together with the total number of migrants and the total number of highly skilled migrants for both 1990 and 2000 are presented in Table A.1. In 16 cases, data are obtained directly from the relevant destination countries' national statistical offices. In 24 cases, data are taken from the IPUMS-International or the United Nations' Economic Commission for Latin America and the Caribbean (CEPAL) databases, two of the largest archives of publicly available census samples. They are based on samples of at least 5 percent to the whole population. Data for the 6 Gulf Cooperation Council (GCC) countries are estimated on the basis of their Labor Force Surveys. Usually, primary data only provide limited details on the country of origin and education level of migrant workers. For example, Saudi Arabia's Labor Force Survey distinguishes a dozen main origin countries while reporting an important residual category. In addition, it only provides the aggregate proportion of post-secondary educated guest workers. We rely on the database provided in OPSW (2010) to split residuals by country of origin and assume that educational structures are homogenous across source countries. The same methodology is applied to the other GCC countries and Saudi Arabia's education breakdown is applied if observations are otherwise missing. We believe these assumptions deliver a reasonable approximation of human capital flows to the Persian Gulf.

Adding 42 destination countries increases the migrant stock by 18.3 million in 2000, including 2.5 high-skilled migrants and 7.3 million women (see Table 1). The proportion of college graduates is 13.4 percent and the share of women is 40.0 percent, far below the ratios observed in OECD destination countries in both of these dimensions (35 percent and 51 percent, respectively). These ratios vary considerably across countries and this heterogeneity is explored in more detail in Section 3. Six of these 42 additional destination countries are home to more than one million foreign-born adults in 2000. These are Ivory Coast (3.9 million), Saudi Arabia (3.1 million), Hong Kong (1.9 million), Israel (1.5 million), the United Arab Emirates (1.2 million) and Malaysia (1.0 million).

2.3 Imputed migration for non-OECD countries

Despite the additional data for 42 non-OECD destinations in 2000 and 27 non-OECD destinations in 1990, our bilateral migration matrices by education levels remain incomplete. One of the key contributions of the paper and the goal of this section therefore, is to predict bilateral migrant stock data by both gender and education for the missing 119 non-OECD countries in 2000 and the 134 countries in 1990. To this end, we first estimate the determinants of the bilateral migrant stocks (disaggregated by gender and education level) for which we have data and then subsequently utilise the resulting parameter estimates from these regressions to predict the gender and education specific bilateral migrant stocks for cells for which we lack data. This section concludes with an analysis of the accuracies of our predictions.

The pseudo-gravity model. The empirical model that we use to make our out-of-sample predictions is based upon recent developments in the literature that are grounded in a theoretical framework based upon an income maximisation approach, which was first introduced into the migration literature by Borjas (1987). The recent innovations within this framework, which ultimately yield estimable pseudo-gravity equations, consider migrants choosing between staying at home or else potentially moving to all other destinations, see for example Beine et al (2011), Hanson and Grogger (2011) and Peri and Ortega (2012). Our model most closely follows the approach of Beine et al (2011). In the model, individuals with different human capital (education) levels choose between alternative destinations and staying at home after observing their individual random shocks. Each country pair (or corridor) is characterized by dyadic migration costs and barriers such as physical distance, linguistic overlap and political linkages. Gender and education specific migration levels are therefore expressed as functions of various bilateral variables in addition to origin and destination specific push and pull factors. The adoption of this model leads to the following specification:

$$M^{jk}_{g,s,t} = \alpha^{g,s}_{0t} + \alpha^{g,s}_{1t} d^{g,s}_{jkt} + \alpha^{g,s}_{2t} A^{g,s}_{kt} + \gamma^{g,s}_{jt} + \epsilon^{g,s}_{ijt} \tag{3}$$

The variables are defined as following:[4] The dependent variable $M^{jk}_{g,s,t}$ is the bilateral migrant stock in the gender-education (i.e. labour force) group (g, s) from country j to country k in the relevant year t - either 1990 or 2000. We have two years and four separate gender-education pairs - high-skilled men, high-skilled women, low-skilled men and low-skilled women. These groupings necesitate the estimation of eight separate regressions in order to determine the relevant year and group specific parameters from which we can make our predictions.

[4]Table A.2 in the Appendix describes the data sources as well as the way we construct and measure these explanatory variables that influence migrant stocks.

The explanatory variables include a set of bilateral geographic distance and linkage (linguistic, political, cultural) variables denoted by $d_{jkt}^{g,s}$. Among these are geographic contiguity, physical distance, colonial relationships and linguistic overlap (see Clair et al., 2004). We also include the total number of migrants by gender from country j in country k at time t based on the bilateral migration database presented in OPSW.[5]

All origin country characteristics (such as economic, political and social push factors) are captured by a set of fixed effects introduced through the vector $\gamma_{jt}^{g,s}$. Unfortunately, we are unable to include an exhaustive set of destination country fixed effects however, since although this would no doubt increase the predictive power of the model, their inclusion would militate against us being able to predict the missing values of interest. Instead, we include a set of ten regional dummies (South Asia, Latin America etc.) as well as various destination specific variables that influence migration patterns, which are denoted by $A_{kt}^{g,s}$. These include whether people in the destination country speak English, the size of the total labor force in the destination country (in logs), GDP per capita (in logs), the total fertility rate (in logs) in the destination country, the ratio of the number of skilled to the total in the destination country and the labor force participation rate in the host nation. A number of dummy variables are also included that capture whether a destination country belongs to the GCC, whether military service is compulsory at destination and whether polygamy is legally practiced. We believe that together with the origin country and destination region dummy variables, these capture the most important determinants of international migration. It is worth noting that all of the destination variables, $A_{kt}^{g,s}$, must be available for all 195 destinations in order to predict the missing migration numbers.

Econometric issues. The presence of a large number of zero or undefined observations in the dependent variables (gender and education specific bilateral migrant stocks) in both 1990 and 2000 give rise to econometric concerns that would yield inconsistent OLS estimates. Zero observations appear in large numbers in many other bilateral contexts such as international trade, official aid, military conflict and polit-

[5]Using alternative data sources might evoke endogeneity problems. Indeed, by definition and apart from measurement errors, our bilateral migration stock aged 25+ is equal to the total stock in OPSW minus migrants aged 0-24. However, our goal is not to identify causal links between variables. By including OPSW stocks, we clearly tolerate endogeneity of some regressors in order to maximize the accuracy and power of our model in predicting the migrant stocks in missing cells in our matrices.

ical alliances. This phenomenon is especially prevalent in migration data sets however, since there is no observed or recorded migration between many country pairs, for example, between the Central African Republic and Peru, due to high geographic, cultural or economic barriers. Furthermore, censuses or alternative surveying instruments are unlikely to capture small migration corridors should a sampling strategy be followed. As a result, we have zero values for about 48.5 percent of the 14,820 observations (195 destination x 76 origin countries) in the aggregate migration matrix for 2000. The ratio of zero observations is 52.6 percent for low-skilled males, 52.9 percent for high-skilled males, 52.8 percent for low-skilled females and 54.0 percent for high-skilled females.[6]

Two main reasons explain why a high proportion of zero observations in the dependent variable typically result in inconsistent paramater estimates. The first is selection bias. Since observations including a zero value in the dependent variable will be dropped from estimation, an inherent selection bias will be introduced since the occurance of zero oberved flows are non-random. In a double log regression model, authors frequently add one to the value of the dependent variable and then take the log, but again this leads to numerous observations being dropped from estimation using standard regression techniques. The second bias has been well documented by Santos Silva and Tenreyro (2006) who demonstrate in their influential paper, in the presence of numerous zeroes in the dependent variable, that the expected value of the error term will be correlated with some of the independent variables should the variance of the error term also be correlated with the independent variables. In other words, in the presence of numerous zero observations in the dependent variable and heteroskadasticity that one of the key assumptions of the OLS model will be violated, namely that the expectation of the error term will be non-zero. In order to surmount both of these issues, Santos Silva and Tenreyro (2006) advocate the use of Pseudo-Poisson Maximum Liklihood (PPML) estimator that yields consistent parameter estimates even in the presence of numerous zero observations in the dependent variable. We therefore deem the PPML estimator as the most appropriate technique for obtaining our parameter estimates. Observations are weighted by the log of the aggregate migrant stock and robust standard errors are always implemented.

[6]Similarly, in 1990, we have zero values for 43.2 percent of the 11,895 observations (195x61 countries) in the aggregate matrix. Similar ratio is 46.9 percent for low-skilled males, 49.0 percent for high-skilled males, 47.2 percent for low-skilled females and 50.9 percent for high-skilled females.

Estimation results. The results for the determinants of gender and education specific migration patterns for 2000 and 1990 are presented in Table 2. Clearly, the OPSW variable is highly significant and explains the largest share of the variation. Its absence causes other variables to become more significant in the estimation, or conversely its inclusion results in some key variables to lose significance. These include per capita income and distance, especially in the estimation of high and low skilled males (columns 3 and 4). The OPSW variable is an excellent predictor for the size of bilateral corridors. Other determinants mainly explain their structure by education and gender.

Colonial relationships encourage all types of migrants equally, while common language has a stronger effect on skilled migrants. Sharing a common border has no effect upon the migration of the highly skilled but encourages greater numbers of the less educated, regardless of gender. Country pairs that are geographically further from one another experience less female migration (due to higher migration costs), although no effect is found in the case of males. While higher fertility levels in destination countries attract all migrants (with a stronger effect on males), prevalence of polygamy at destination deters all migrants, high-skilled females being most affected. Compulsory military service in the destination country however, only deters the migration of the less educated.

As far as education is concerned, we find evidence of migrant selection. Countries with a higher proportion of highly skilled workers attract higher proportions of highly skilled migrants. The oil-rich GCC (Gulf Cooperation Council) countries also attract larger numbers of highly skilled as do countries with larger labor forces, although this also deters less educated females. The existing literature (see Grogger and Hanson, 2011, and Beine et al., 2011), shows that the educational composition of migrants from more distant countries is biased towards the more highly skilled. This would be in accordance with our findings from the contiguity variable, but this is not confirmed by our distance variable.

The results for 1990, broadly reflect the same patterns exhibited in the 2000 data, although for this decade the OPSW variable exerts even more influence over the remaining variables. The effects of a common border, a common language, larger bilateral migrant stocks and higher fertility rates all broadly have the same effect in 1990. While the distance variable is largely insignificant, per capita GDP has a positive effect on less educated workers. Sharing a colonial link and greater degrees of labor force participation have positive effects on both skilled males and females but no effect on the less educated. In 1990, the prevalence of legal polygamy only effects women, especially the highly skilled.

[INSERT TABLE 2 AROUND HERE]

Accuracy of the in-sample predictions. The main goal of this section is to predict the aggregate number of high and low skilled males and females in the destination countries for which our bilateral data are missing. Our predictions are based on the four gender and skill specific estimations presented in the previous section. In order to gain confidence, we perform several validation exercises. The predicted values of the migrant stocks (195x195 matrix for each decade) comprise two parts. The "in-sample" predictions, correspond to the migrant stocks in the 76 destination countries in 2000 and 61 countries in 1990 for which we have collected data. The "out-of-sample" predictions instead refer to the remaining bilateral migrant stocks for which raw data are missing. Comparisons of the in-sample predictions to the actual values therefore, potentially yield many insights as to the accuracy and efficiency of our estimates.

Table 3 provides summary statistics of our in-sample predictions for the four groups of migrants in the year 2000.[7] In each section of Table 3, the first column provides the number of observations. The second column *'Corr'* provides the correlation between actual and predicted values. The third column *'Migs'*, gives the total number of actual migrants (in millions,) while the fourth provides the migrants as a percentage of the relevant gender and education specific sample. The final two columns provide the mean averages and the standard deviations of the logarithm of the ratio of our predicted values to the actual numbers. These are the same things as the differences of the logs and arguably provide the best summary statistics of exactly how accurate our predictions are.[8] These statistics are compiled for different sizes of migration corridors. The first row is for all observations; the second, third and fourth rows are when the predicted values and the actual values are strictly less than 250, when the predicted and actual values are greater than 250 and when both are greater than 5000. The fifth and sixth rows are for OECD and non-OECD destinations when the predicted and actual values are both greater than 250.

[7] Very similar patterns are obtained for 1990 and are available upon request.

[8] Note $\ln(a/b) = \ln(a) - \ln(b)$. We use the log ratio for comparison to treat over and under predictions symmetrically when the averages are taken.

Clearly, the correlations are high overall (row 1), indicating *prima facie* that the PPML estimator perform well. However, there is significant variation when we focus upon small corridors. For example, in 2000, the correlation between the predicted and the actual stocks is around 0.60, if the predicted and actual migrant stocks are both less than 250. This correlation rises to 0.99 when the stocks are greater than 250. This variation is simply due to measurement errors and the strong influence of unobservable factors on smaller corridors. Note that the corridors with fewer than 250 migrants account for less than 1 percent of the total number of migrants even though there are a very large number of them. This is closely related to the prevalence of large number of zeros in the migration matrices where a small number of large corridors account for the vast majority of migrants. In other words, although prediction biases are clearly important for a large number of corridors, these tend to be small and relatively immaterial to the analysis of global migration patterns. Our estimates of the more important large corridors and, therefore, our analysis of the global migration patterns are not overly distorted. As expected, the correlations are stronger for the OECD countries in our sample when compared to non-OECD countries reflecting superior quality data.

The correlations between the predicted and the actual migrant stocks only capture part of the story. The next exercise analyzes the log of the ratio of the predicted stocks to the actual stocks of migrants. If our predictions were flawless, all of these ratios would be equal to 0. In Table 3, columns *'Mean LR'* and *'Sd LR'* give the unweighted means and standard deviations of these log ratios by corridor size and higher values indicate lower precision. The most striking numbers are in the first and second rows. The mean of the log ratios range from 0.18 (for low-skilled males) to 0.32 (high skilled females), implying significant deviations of the predicted values from the actual ones. However, when we focus upon larger corridors, the summary statistics of these log ratios improve dramatically. For example, for corridors above 250 (row 3), the mean of the log ratio ranges between 0.05 to 0.10. Since these groups comprise over 99 percent of our in-sample migration stock, we conclude that our estimates imply relatively small overall global biases.

[INSERT TABLE 3 AROUND HERE]

Figure 1 provides additional insights about the distribution of the log of the ratio of predicted stocks to actual stocks. Log ratios are used, instead of simple ratios, to reflect upward and downward biases symmetrically. We focus upon high-skilled females, but the patterns for the other gender and skill groups (and for 1990) are identical.

Each figure presents graphically the distribution of the (log) ratios on the real line. A distribution more closely centred around 0 therefore implies more precise estimation.[9] We present the results for all corridors, less than 250, greater than 250 and greater than 5000 (first four rows in Table 3). These figures highlight the large inherent idiosyncratic factors that exist in the estimation of small migration corridors. As clearly seen in the comparison of Figures 1b, 1c and 1d, the distributions of small corridors cover a much wider range indicating a larger standard deviation. Conversely, as the corridors become larger, the distribution of the log ratio becomes more concentrated around zero and symmetrical. This demonstrates the increasing precision when predicting larger migration corridors.

[INSERT FIGURE 2 AROUND HERE]

2.4 Labor force data by education level

Although often informative, migrant stocks are not the right measures to compare the countries of different sizes and development levels. In order to accurately quantify the intensity of emigration (or immigration), emigrant (or immigrant) stocks need be expressed relative to an appropriate measure of the relevant labor force. In this paper we argue that the most appropriate measure is that of the *natural* labour force, i.e. the number of workers from a particular origin country regardless of where they currently reside. The ability to recover our measure of the natural labour force $N^i_{g,s,t}$, a prerequisite for which is to have measures of immigrant/emigrant stocks for all nations in the world, is a key contribution of the current work since it allows a more nuanced understanding of the mobility of human capital internationally. Given our estimates of immigration and emigration globally, before we can estimate $N^i_{g,s,t}$ using Eq. (2), we first need to construct a consistent measure of $L^i_{g,s,t}$, i.e. the resident labour force.

We begin with a measure of the total working-age population (i.e. aged 25 and

[9] If the actual and predicted stocks are equal to each other, we would have $\ln(predicted/actual) = \ln(1) = 0$

over) by gender as provided by the United Nations.[10] Data are missing for a few countries and these are instead obtained from the CIA world factbook.[11] These data are then split across skill (i.e. education) groups using international indicators of educational attainment. Here, we follow Docquier and Marfouk (2006) or Docquier, Lowell and Marfouk (2009) and combine different data sets documenting the proportion of post-secondary educated workers in the population aged 25 and over (i.e. De La Fuente and Domenech, 2006, Barro and Lee, 2001, and Cohen and Soto, 2007). The post-secondary concept corresponds to a broad definition of high-skill labor as it includes workers with at least one year of college or university. This definition is relevant for developing countries, where the share of college graduates in the labor force is lower than one percent in some cases. Given the construction of $L^i_{g,s,t}$, Eq. (2) is then used to identify the size and structure of the natural labor force, $N^i_{g,s,t}$, for each labor type, country and period.

2.5 Improved brain drain indicators

Our dataset enables us to improve the skilled migration indicators presented in previous studies (Docquier and Marfouk, 2006, Dumont and Lemaitre, 2004, Docquier, Lowell and Marfouk, 2009 or Dumont, Martin and Spielvogel, 2008). These works provide cross-country data on the relative intensity of emigration (referred to as emigration rates), controlling for the population size and the skill structure in the origin country, while focusing upon the subset of OECD destinations.[12] Such measures necessarily omit emigrants that reside in non-OECD destinations therefore, which in turn leads to biases that are especially severe for countries that send a large proportion of their emigrants to non-OECD nations.

Instead, our refined gross emigration rates ($e^i_{k,s,t}$) and net emigration rates ($b^i_{k,s,t}$) for a given country i are defined as follows :

$$e^i_{g,s,t} \equiv \frac{E^i_{g,s,t}}{N^i_{g,s,t}}, \quad b^i_{g,s,t} \equiv \frac{E^i_{g,s,t} - I^i_{g,s,t}}{N^i_{g,s,t}} \tag{4}$$

[10] Population data by age and gender are provided by the United Nations Population Division and can be found at http://esa.un.org/unpp.

[11] See http://www.cia.gov/cia/publications/factbook.

[12] Data on selected non-OECD destination countries were included in Docquier and Rapoport (2011) and in the latest version of the OECD database.

so that (2) can be written as $L^i_{g,s,t} \equiv N^i_{g,s,t}(1 - b^i_{g,s,t})$.

In comparison to the existing literature, the current paper contributes three major improvements to the measurement of international human capital mobility:

- *Comprehensiveness* – Existing studies record immigrants in a limited set of destination countries (OECD countries in addition to a few selected non-OECD destinations). By expanding the number of destinations to cover all countries in the world, we provide a comprehensive picture of international human capital mobility. Furthermore, we are able to calculate total emigrant stocks, $E_{g,s,t}$ for all the countries of the world, since we present comprehensive migration matrices. For example, compared to the set of OECD destinations, the total number of adult migrants identified in 2000 increases from 59.3 to 100.5 million.

- *Natural-based* – We are able to refine our definition of emigration rates. Instead of dividing the number of migrants by the corresponding labor force at origin (which includes immigrants), we divide it by the natural labor force, i.e. the number of individuals born in the origin country (which excludes immigrants). Our emigration rates thus differ from those computed in previous studies, $E_{g,s,t}/(L_{g,s,t} + E_{g,s,t})$ as we do not need to proxy the natural labor force $N_{g,s,t}$ at the denominator with $L_{g,s,t} + E_{g,s,t}$. This makes a substantial difference in countries with large levels of immigration, especially for hig skill levels.

- *Net vs Gross* – We are also able to identify the size and skill structure of adult immigration in all host countries, including the developing world. Those immigration data were only available for OECD member states and selected non-OECD countries in previous works. Hence, we can compare entries and exits of workers and compute comparable net migration balances for college graduates and less educated workers for all nation states.

Given the scope of our dataset, are are also able to characterize the skill level of the natural population. Given (2) and (4), average skill levels of naturals and residents are linked by the following equation where the subscripts h and l stands for the high and low skilled, respectively:

$$\frac{L^i_{k,h,t}}{L^i_{k,l,t}} \equiv \frac{1 - b^i_{k,h,t}}{1 - b^i_{k,l,t}} \cdot \frac{N^i_{k,h,t}}{N^i_{k,l,t}}.$$

As can be seen, international migration affects average human capital levels if emigrants and immigrants differ from non-migrants in terms of their skill composition or if net emigration rates differ across skill groups ($b^i_{k,h,t} \neq b^i_{k,l,t}$). Many studies have documented and explained the pattern positive selection in international migration ($e^i_{h,s,t} > e^i_{l,s,t}$). However, what matters is the level of the net emigration rates of high-skilled and low-skilled workers: international migration reinforces human capital inequalities across nations if $b^i_{k,h,t} > b^i_{k,l,t}$. We will illustrate this phenomenon by comparing the concepts of human capital per natural and per resident, measured by the following indicators:

$$H^i_{g,t} = \frac{N^i_{g,h,t}}{N^i_{g,l,t} + N^i_{g,h,t}}; \quad h^i_{g,t} = \frac{L^i_{g,h,t}}{L^i_{g,l,t} + L^i_{g,h,t}} \quad \forall g, t$$

where $H^i_{g,t}$ is the proportion of college graduates among naturals of gender g, and $h^i_{g,t}$ is the same proportion computed on the resident labor force.

3 Results

Combining detailed original raw data with predicted (imputed) numbers, we construct a global dataset comprising gender and skill specific migrant stocks for 195x195 corridors in 1990 and 2000. Our final correction is to set the migrant stocks to zero if the predicted (gender and skill specific) migrant stock in a given corridor is below 250. As explained above, this is due to the lack of precision in the estimation of small corridors which account for a relatively large set of observations but less than one percent of the overall global migrant stock.

In this section, we provide some general statistics and highlight the global patterns of international migration in 1990 and 2000 (Section 3.1). We then study country-specific characteristics and identify the main source countries, focusing on college-graduate migrants (Section 3.2), and college-graduate female migrants (Section 3.3). Finally, we will compare the concepts of human capital per resident and per natural (Section 3.4).

3.1 Global patterns

Table 4 details total emigrant stocks and their education/gender composition in 2000 and 1990 for key regions or income categories of the world. The top portion of Table 4 isolates the group of OECD countries and divides the world into high-income and developing countries. We then distinguish between low income, least developed and small island developing states (SIDS) which have unique migration patterns. The second section of the table divides the world into twelve geographical regions: (1) the United States, (2) Canada, Australia and New Zealand as a single entity, which is referred to as CANZ, (3) the twenty-seven nations of the European Union (EU27), (4) the oil rich Gulf Cooperation Council (GCC) countries, (5) Latin America and the Caribbean (LAC), (6) Sub-Saharan Africa (SSA), (7) the countries of the Commonwealth of Independent States (CIS), (8) India, (9) China, and (10) countries in the Middle East and North Africa excluding the GCC (MENA). We do not report results for the heterogeneous set of remaining countries.

Beginning in the top-most panel, the figures reveal that as income levels increase so do the proportions of both the percentages of high-skilled emigrants and female emigrants abroad. Comparing emigrations from these regional groupings to OECD and non-OECD destinations further reveals the strong selection inherent in world migration patterns. Across all regional groups, a far higher proportion of both college educated and women emigrate to OECD destinations. This selection on skills is most pronounced in the cases of low income and least developed countries from which only 3.7% of emigrants to non-OECD nations have college education as opposed to 38% and 34.6% in OECD nations respectively. These patterns are also reflected strongly in the data for 1990. Interestingly, the only region to send more female emigrants to non-OECD destinations is the grouping of small island developing states.

The second sections of the top and bottom panels of Table 4, again reveal strong patterns of selection, since the proportions of both the high-skilled and women emigrants are far larger in OECD destinations when compared to non-OECD destinations; with the exception of women from Latin America and the Caribbean who have a greater tendency to emigrate to non-OECD destinations. This almost certainly reflects intra-regional migration in that part of the world. Examining how this selection between OECD and non-OECD destinations has changed over time - in other words the difference of the differences - also yields interesting results. The selection of emigrants from all regions to OECD nations, in terms of high-skill composition increased between 1990 and 2000, with the exception of emigrants from the Gulf Cooperation Council and Latin America and the Caribbean. In the case of emigrants from India and the countries of the Commonwealth of Independent States this increase in selection (between OECD and non-OECD countries) of the high-skilled has increased by 11% and 16% respectively. An examination of how the selection of female migrants between OECD and non-OECD destinations over time, interestingly reveals the opposite pattern however. Although, as already discussed, most countries exhibit migrant selection in terms of the proportions of females they send to OECD as opposed to non-OECD countries, the extent of this selection decreased between 1990 and 2000, implying that non-OECD nations are increasingly important destinations for female emigrants. For the countries of the Gulf Cooperation Council there was no difference in this selection over time, while for emigrants from China and sub-Saharan Africa the selection of female migrants has reversed. In 1990 for both these groups proportionally more female migrants emigrated to non-OECD nations, while in 2000 females from both groups instead favoured OECD destinations.

[INSERT TABLE 4 AROUND HERE]

Columns 1 and 4 in Table 5 provide gross and net emigration rates, calculated according to Eq. (4). For gross rates, we further distinguish between emigration to OECD and non-OECD countries (columns 2 and 3). Net rates are provided for men and women with college education (columns 5 and 6). Globally, gross high-skilled emigration rates decrease with country size and income level, which is a finding in accordance with the previous literature. The groups of small developing islands and least developed countries are most affected, with high-skilled emigration rates of 40.9 and 20.2 percent, respectively. The most affected geographic regions are the MENA (17.6 percent), sub-Saharan Africa (15.0 percent) and Latin America and the Caribbean (12.1 percent). The role of non-OECD destinations varies across groups. High-skilled emigration to non-OECD countries is negligible for high-income and small islands developing states. Conversely however, high-skilled emigration to non-OECD countries accounts for about one third of the brain drain from lower-income countries and is of particular significance for the countries of MENA, the ex-Soviet block, the GCC as well as India.

A comparison of gross and net emigration rates proves highly instructive. High-income and OECD countries exhibit negative net high skilled migration rates meaning that the incoming pool of educated people to those regions more than compensates for any human capital loss suffered as a consequence of their skilled nationals emigrating abroad. Consequently, international high-skilled mobility increases the number of college graduate workers in the labor force by over 10 percent in the United States, around 30 percent in other settlement countries (Canada, Australia and New Zealand) and remarkably doubles this proportion in oil producing countries. With regards developing regions, gross and net rates are strongly correlated although net rates are sensibly lower. Another advantage of calculating net migration rates at the regional level is that they remove intra-regional movements. This explains why net brain drain rates are much lower than gross rates in the MENA and CIS regions, two regions characterized by large internal migration flows. Turning finally to gender differences, the final columns of Table 5, demonstrate that in all regions net emigration rates are lower for males than for females, with the exception of the EU27 and MENA.

[INSERT TABLE 5 AROUND HERE]

3.2 Country-specific results

Our exploration of the impact of skill transfer around the globe highlights the importance of our introducing non-OECD destinations into our analysis. Collectively their introduction serves to highlight significant heterogeneity across countries and within regions. The aim of this section is to present some important and insightful country-specific stylized facts. Figure 3 illustrates the effect of introducing non-OECD countries into our analysis upon the distribution of high-skilled emigration rates. Although the average share of non-OECD destination in high-skilled migration is around 20 percent (5.3 million over 26.2 in 2000 and 3.8 over 16.4 in 1990, as shown in Table 1), the variance of this share is large. Figure 3.a plots the distribution of the ratio of non-OECD to total gross emigration rates in 2000 for college graduates.[13] The peak of this kernel density corresponds to a ratio of just 0.065 and in the majority of cases (123 out of 195), the ratio does not exceed 0.20. However the distribution is heavily right-skewed such that this ratio exceeds 0.50 in 32 countries, i.e. in no less than a sixth of the sample. The individual countries that comprise the thick right-hand tail of the distribution include nations of the Middle-East (that predominantly send emigrants to oil producing countries), Southern African nations

[13] We use the gaussian kernel density estimator implemented in Stata.

(that principally send migrants to the Republic of South Africa) and ex-Soviet-block members, which are characterized by significant and voluminous migrations between one another.

Unsurprisingly, for many countries, a significant disparity exists when comparing high-skilled emigration rates to all destinations and to the OECD alone, which until now has been the focal group in the literature. These marked differences are illustrated in Figure 3.b, which plots the gross emigration rates of college graduates to OECD destinations on the x-axis, against those to all destinations on the y-axis, for each country. In doing so, the figure highlights the importance of our comprehensive global approach. Each bubble in Figure 3.b represents an origin country and the size of the bubble is proportional to the high-skilled emigration stock from that country.

Overall there exists a strong correlation between our (OECD-) restricted and our global measures; but in many cases, the inclusion of non-OECD destinations has a dramatic impact on the magnitude of our estimates of high-skilled emigration rates, i.e. the gross brain drain, for many poorer developing countries. Examples of these differences include a 53 percentage point difference for West Bank and Gaza, 37 for Yemen, 28 for Palau, 27 for Namibia and 25 for Jordan, which tend to send emigrants to other countries in their regions. Changes are significantly lower for the small islands of the Pacific and the Caribbean where the largest emigration rates are observed since these countries predominantly send emigrants to North America, Australia and New Zealand.

Figure 3.c compares the gross and net emigration rates of college graduates on the horizontal and vertical axes, respectively, shows the advantage of using net rather than gross rates. Obviously, net rates (exits minus entries) are by definition lower than gross rates (exits) so that the whole scatterplot lies beneath the 45 degree line. Net rates are high and similar to gross rates in small island developing states, but they are negative in high-income countries and, especially, in the GCC countries.

[INSERT FIGURE 3 AROUND HERE]

Finally, Table 6 lists the 25 countries with the highest (left panel) and lowest (right panel) net emigration rates of college graduates, excluding small states with less than one million workers (population aged 25+). Nine sub-Saharan African countries belong to the top-25. Other remarkable cases are Jamaica (84.7) percent) and Haiti (80.7 percent), Laos (44 percent), Afghanistan (44.8 percent) and four other countries that lose more than 30 percent of their college educated labor force (Bosnia and Herzegovina, Yemen, El Salvador and Sri-Lanka). Among the main net receivers, we find many high-income OECD and oil producing countries but also countries such as Ivory Coast, Paraguay and Papua New Guinea where relatively few natives have college education.

[INSERT TABLE 6 AROUND HERE]

3.3 Female Brain Drain

Female brain drain is matter of deep concern, not least since it is recognized that women's human capital is an important determinant of labor productivity, children's education and economic growth (see for instance Knowles et al., 2002, Coulombe and Tremblay, 2006, Klasen 2000, Dollar and Gatti, 1999, Blackden et al., 2006). Societies that are characterized by a failure to invest in female education or else those that lose a high proportion of educated women through emigration are therefore likely to exhibit slower growth rates and lower income levels. Conversely, societies that experience a net female skill gain may experience higher growth rates.

Figure 4 illustrates graphically the impact of our introducing non-OECD destinations into our analysis of female high-skilled emigration rates. In Figure 4.a, we compare the high-skilled emigration of men (x-axis) and women (y-axis). Most observations (141 out of 195) lie above the 45 degree line, indicating that brain drain is more pronounced in the case of females (when compared to males). On average, the brain drain for females is 15 percent higher than for males, as illustrated on Figure 4.a by the linear trend estimated for the whole sample.[14] Such gender disparities are particularly apparent from sub-Saharan African countries and more broadly in cases in which women have poorer access to human capital. The intensity of college-educated women emigration is greater to OECD destinations however, such that the inclusion of non-OECD destinations has less bearing on our analysis of female brain drain in comparison with the impact on total high-skilled emigration rates, as demonstrated by comparing Figure 4.b and Figure 3.b. Nevertheless, the ratio of non-OECD to total female gross emigration rates in 2000 exceeds 0.50 in 25 countries (as opposed to the 32 cases taking men and women together). Similarly to Figure 3.c, Figure 4.c

[14]Focusing on OECD destination countries, the gap increases to 18 percent (see Docquier, Lowell and Marfouk, 2009).

plots gross and net emigration rates of college graduates only this time focusing solely upon female migration. Although as previously noted, the impact upon our analysis of high skilled female migration is less pronounced when we introduce non-OECD destinations, Figure 4.c nevertheless highlights the fact that wealthier countries gain relative to poorer nations since they are more successful in attracting higher volumes of college educated females.

[INSERT FIGURE 4 AROUND HERE]

Table 7 lists the 25 countries with the highest (left panel) and lowest (right panel) net emigration rates of female college graduates, excluding small states with less than one million workers (population aged 25+). By and large the entries are similar to those in Table 6 although the magnitude of the net losses are broadly larger for the most affected countries. New entries in the left panel include the West Bank and Gaza and several African nations, namely, Cameroon, the Democratic Republic of Congo, Nigeria and Rwanda; meaning that for these countries the magnitude of skilled emigration rates are particularly skewed in favour of women. The only new entries in the right panel in Table 7 (as when compared to Table 6), include Nepal, Turkmenistan and Japan, meaning that these destinations are particularly attractive to college-educated female migrants relatively to their natural female population of college graduates.

[INSERT TABLE 7 AROUND HERE]

3.4 Brain drain and human capital

Our final piece of analysis draws upon the recent contribution of Clemens and Pritchett (2008), who provide comparable measures of income based upon the concept of the natural population. They argue "If economic development is that which raises human well-being, then crossing international borders is not an alternative to economic development; it is a form of economic development". They estimate income per natural, the mean annual income of persons born in a given country regardless of where that person resides and compare it with the standard indicator of income per capita, based upon a specific geographic area. Since human capital mobility affects both incomes per natural and the more usual measure of income per capita, it is instructive to compare measurements of high-skilled migration for both the resident and the natural population.

In line with our earlier expression for the average skill levels of naturals and residents, Figure 5.a graphically compares the high-skilled emigration rates of natives (i.e. naturals) on the x-axis and of residents on the y-axis. Since most countries that deviate from the 45 degree line lie beneath it, this shows that in general, countries' natural work force is more highly educated than the workforce that resides in that country ($h^i_{w+m,t} < H^i_{w+m,t}$). In other words, high-skilled immigration to these nations fails to compensate for the skill losses endured when college-educated natives move abroad or else that those countries characterized by net entries of college graduates also experience greater net inflows of less educated workers. For the year 2000, we identify 42 cases with negative net high-skilled emigration rates, but globally migration only increases human capital in 27 of them. In the remaining 24 countries, net entries of college graduates fail to compensate for net inflows of less educated workers; this group includes developing countries such as Cote d'Ivoire, Costa Rica, Gabon and Russia, but also wealthier countries such as Belgium, France, Germany, Luxembourg, the Netherlands, New Zealand, Norway and the United States among others. Similarly, Figure 5.b repeats the previous exercise only this time restricting the analysis to females, the results from which are broadly similar. For the year 2000, we identify 36 cases with negative net high-skilled emigration rates, but global migration only increases women's human capital in 26 of them.

[INSERT FIGURE 5 AROUND HERE]

4 Conclusion

The absence of harmonized, detailed and reliable data on bilateral migration patterns by skill and gender has restrained proper analysis of many important economic, social and political issues relevant for economic growth, development, poverty and income distribution. In this paper, our goal is to take an important step to address these constraints and construct a comprehensive database documenting bilateral movements of high and low skilled workers by gender for 1990 and 2000. In addition to identifying important patterns, especially on skilled migration from poorer developing countries, we hope to provide a readily available dataset for other researchers. We use recorded bilateral migration data for a group of countries to estimate the determinants of migration which are then used to predict the missing data for many destination countries. We find that, even though we lack raw data for a large number of destination countries, they account for a relatively small portion of the global migration stocks and our estimates are quite reliable. After we evaluate the precision of our predictions, we combine the recorded data with the most reliable estimates to construct comprehensive 195x195 migration matrices.

Even the simple tables and figures we produce from our full migration matrices reveal that focusing only on migration to OECD countries might hide important patterns. We see that migration to non-OECD countries account for 40% of world migration and 20% of high-skilled migration. Even though migration to non-OECD destinations is less skill intensive, it is very high for low income and least developed countries. In other words, a narrow focus on OECD destinations prevents us from assesing the full extent of skilled migration from many countries that suffer the most from it. These gaps are even wider for many individual countries and more detailed analysis will naturally lead to more interesting conclusions. The gender dimension of skilled migration from developing countries is also an important issue that deserves further scrutiny. Comparison of 1990 and 2000 data reveal rapidly changing patterns, especially as women increase their education levels but cannot seem to find appropraite jobs in their local markets and choose to emigrate.

This database is clearly an evolving product that can be progressively improved by replacing estimates with official data or adding new census rounds. In addition, we do not control for quality of the education but rather take a university degree, regardless of the location it was obtained, at face value. The next rounds, as the OECD database does, can attempt to control for the human capital content of the education level via various measures such as occupational placement. While we quantify migration stocks by region and income group, identify the main sources and destinations of human capital and the most important migration corridors, the database will prove to be key in understanding the demographic, economic. sociological and political implications of international migration for scholars of many different fields.

5 References

Barro, R., J and J-W., Lee (2001), "International Data on Educational Attainment: Updates and Implications," Oxford Economic Papers, vol. 53(3), pages 541-63.

Beine, M., F. Docquier and C. Ozden (2011), "Diasporas", Journal of Development Economics, 95(1), pages 30-41.

Blackden, M., Canagarajah, S., Klasen, S., Lawson, D., (2006), "Gender and growth in Sub-Saharan Africa: Issues and evidence", UNU-WIDER Research Paper no. 2006/37, UNUWIDER: Helsinki.

Borjas, G. J., (1987), "Self-Selection and the Earnings of Immigrants", American Economic Review, vol. 77(4), pages 531-53.

Clair, G., G. Gaullier, Th. Mayer and S. Zignago (2004): A note on CEPII's distances measures. Explanatory note, CEPII, Paris.

Clemens M. A., and L. Pritchett, (2008), "Income per Natural: Measuring Development for People Rather Than Places," Population and Development Review, vol. 34(3), pages 395-434.

Cohen, D. and M., Soto, (2007), "Growth and human capital: good data, good results," Journal of Economic Growth, vol. 12(1), pages 51-76.

Coulombe, S. and J-F. Tremblay (2006), "Literacy and Growth", Topics in Macroeconomics, vol. 6 (2).

Docquier, F., and A. Marfouk (2004): Measuring the International Mobility of Skilled Workers - Release 1.0. Policy Research Working Paper, no. 3382.World Bank,Washington, DC.

Docquier, F., and A. Marfouk (2006): International migration by educational attainment (1990-2000). In C. Ozden and M. Schiff (eds): International Migration, Remittances and Development. Palgrave Macmillan: New York.

Docquier, F., B.L. Lowell and A. Marfouk (2009): A gendered assessment of highly skilled emigration. Population and Development Review, vol. 35 (2), pages 297-321.

Dollar D and R., Gatti (1999), "Gender inequality, income, and growth: are good times good for women?", World Bank Policy Research Report on Gender and Development, Working Paper Series No. 1.

Doménech R. and A., de la Fuente (2006), "Human Capital in Growth Regressions: How much difference does data quality make?", Journal of the European Economic Association, vol. 4(1), pages 1–36.

Dumont, J.C., and G. Lemaître (2004): Counting immigrants and expatriates in OECD countries: a new perspective. Mimeo: OECD.

Dumont, J.C., J.P. Martin and G. Spielvogel (2007): Women on the move: the neglected gender dimension of the brain drain. IZA Discussion Paper 2920.

Grogger, J., and G.H. Hanson (2011): Income maximisation and the selection and sorting of international migrants. Journal of Development Economics, vol. 95(1), pages 42-57.

Klasen, S., (2000), "Does Gender Inequality Reduce Growth and Development? Evidence from Cross-Country Regressions", Collaborative Research Center 386, Discussion Paper 212 .

Knowles, S., P. K., Lorgelly and P. D. Owen (2002), "Are educational gender gaps a brake on economic development? Some cross-country empirical evidence", Oxford Economic Papers, vol. 54(1), pages 118-149.

Ozden, C., C.R. Parsons, M. Schiff and T. Walmsley (2011): Where on earth is everybody? The evolution of global bilateral migration 1960-2000? World Bank Economic Review vol. 25(1), pages 12-56.

Parsons, C.R., R. Skeldon, T.L. Walmsley and L.A. Winters (2007): Quantifying international migration : a database of bilateral migrant stocks. Policy Research Working Paper 4165: World Bank.

Peri, G. and F. Ortega (2012), The Effect of Income and Immigration Policies on International Migration, NBER Working Paper No. 18322

Santos Silva, J.M.C., and S. Tenreyro (2006): The log of gravity. Review of Economics and Statistics, vol. 88(4), pages 641–658.

6 Appendix

6.1 Data sources

The statistical appendix in DLM (2009, p. 317) describes the data sources for the 30 OECD countries in their sample. Table A1 below describes the various data sources used for the 46 additional countries covered in this study. It provides the total number of migrants together with the number of high-skill immigrants recorded in 1990 and 2000, by destination country. Country order is governed by the size of the total immigration stock in 2000.

[INSERT TABLE A1 ABOUT HERE]

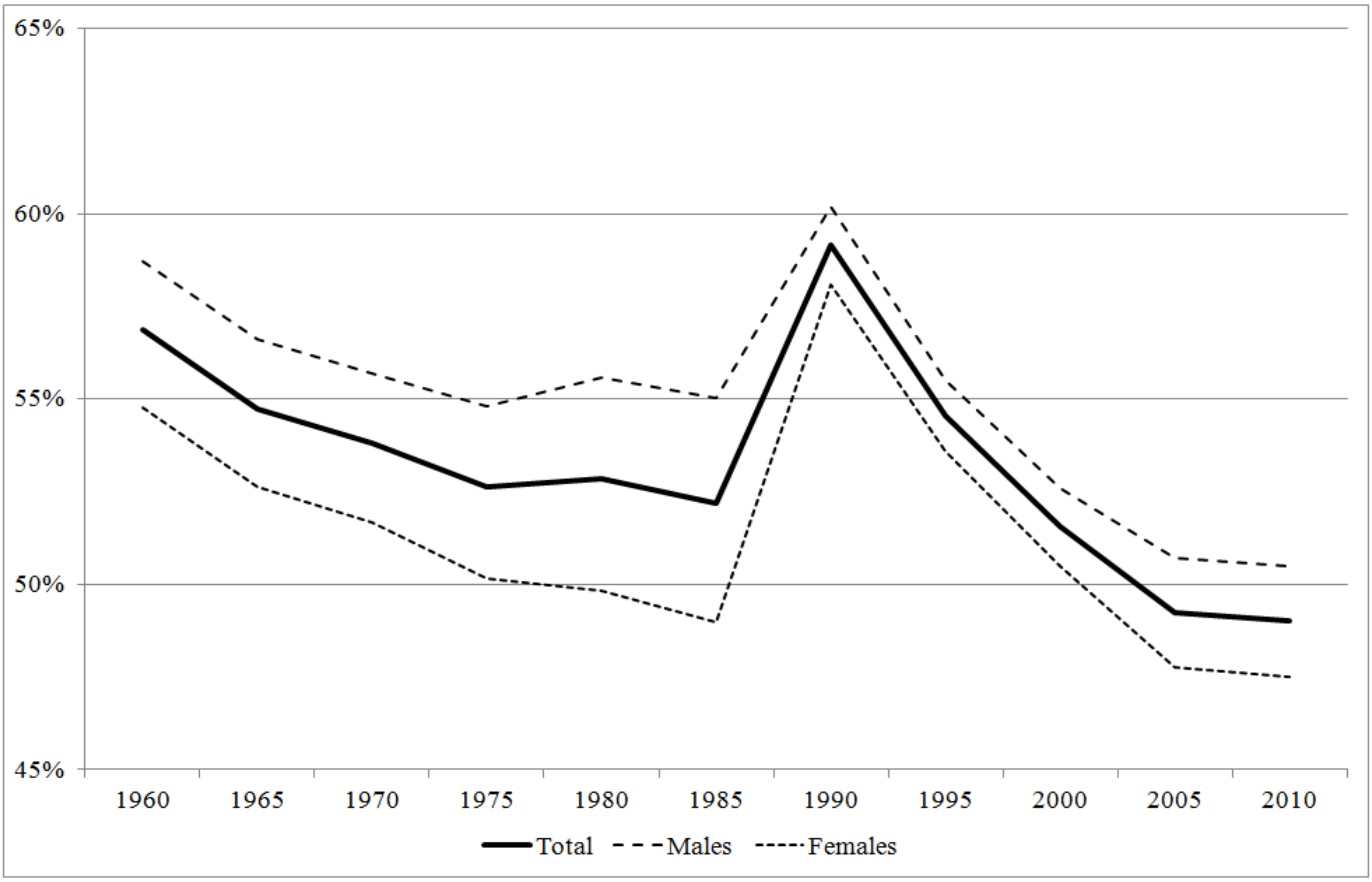
6.2 Explanatory variables

Table A2 describes the data sources for the explanatory variables used in regressions of Section 2.3.

[INSERT TABLE A2 ABOUT HERE]

TABLES AND FIGURES

Figure 1. Share of non-OECD destinations in the world migration stock (data by gender, 1960-2010)



Source: United Nations Population Division (2007, 2012)

Table 1. Migration stocks 25+ in 1990 and 2000 (in millions)

	Total (million)	**To OECD**[a] (million)	**To non-OECD**[a] (million)	(%)[b]	**Including imputed stocks** (milion)	(%)[b]
Year 2000						
Total	100.5	59.3	41.2	41.0	22.9	22.8
College graduates	26.2	20.9	5.3	20.1	2.8	10.7
Less educated	74.3	38.4	36.0	48.4	20.1	27.1
Males	51.5	29.0	22.5	43.6	11.5	22.3
College graduates	13.8	10.6	3.2	23.2	1.6	11.5
Less educated	37.7	18.5	19.3	51.1	9.9	26.2
Females	49.0	30.3	18.8	38.3	11.5	23.4
College graduates	12.4	10.3	2.0	16.5	1.2	9.8
Less educated	36.6	19.9	16.7	45.7	10.2	28.0
Year 1990						
Total	80.2	42.6	37.7	47.0	25.4	31.7
College graduates	16.4	12.6	3.8	23.1	2.5	15.3
Less educated	63.9	30.0	33.9	53.1	22.9	35.9
Males	42.2	21.0	21.2	50.3	13.5	31.9
College graduates	9.2	6.7	2.6	27.6	1.6	17.8
Less educated	33.0	14.3	18.7	56.7	11.8	35.9
Females	38.0	21.6	16.4	43.2	11.9	31.4
College graduates	7.1	5.9	1.2	17.4	0.9	12.0
Less educated	30.9	15.7	15.2	49.2	11.1	35.9

Notes. [a] 34 OECD destination countries; [b] Share of migrants to non-OECD countries, and imputed migration stock, in total migration.

Table 2. Poisson regressions (dependent = migration stock by gender and education)

	2000				1990			
	High-skill fem	Low-skill fem	High-skill males	Low-skill males	High-skill fem	Low-skill fem	High-skill males	Low-skill males
Common border	-0.031	0.247**	0.091	0.243**	0.051	0.308**	0.020	0.235**
	(0.102)	(0.114)	(0.107)	(0.100)	(0.136)	(0.127)	(0.130)	(0.112)
Distance	-0.133***	0.100**	-0.009	-0.052	-0.089**	-0.034	-0.006	-0.036
	(0.035)	(0.040)	(0.035)	(0.041)	(0.038)	(0.045)	(0.039)	(0.052)
Common language	0.415***	0.151*	0.400***	0.013	0.197**	0.091	0.219***	-0.045
	(0.070)	(0.088)	(0.063)	(0.076)	(0.092)	(0.107)	(0.071)	(0.096)
Former colony	0.273***	0.234**	0.289***	0.295***	0.244**	0.060	0.199*	0.103
	(0.083)	(0.105)	(0.093)	(0.102)	(0.107)	(0.137)	(0.110)	(0.128)
OPSW bilateral stock (log) [b]	0.684***	0.892***	0.726***	0.932***	0.718***	0.894***	0.739***	0.906***
	(0.034)	(0.024)	(0.026)	(0.202)	(0.020)	(0.026)	(0.019)	(0.025)
Some English [a]	0.318***	-0.020	0.073	-0.202***	0.206**	0.110	-0.138	-0.149*
	(0.034)	(0.090)	(0.080)	(0.064)	(0.087)	(0.095)	(0.085)	(0.084)
GDP per capita (log) [a]	0.175*	0.101	0.051	0.098	0.059	0.278**	-0.263**	0.273**
	(0.105)	(0.099)	(0.099)	(0.099)	(0.140)	(0.127)	(0.127)	(0.132)
Total fertility (log) [a]	0.355**	0.369**	0.564***	0.571***	0.139	0.639***	0.774***	1.007***
	(0.152)	(0.162)	(0.139)	(0.130)	(0.202)	(0.231)	(0.208)	(0.219)
Skill destination workforce [a,b]	1.748***	0.339	2.633***	0.256	0.868	0.008	4.247***	0.269
	(0.469)	(0.585)	(0.519)	(0.515)	(0.669)	(0.637)	(0.755)	(0.610)
Total labor force (log) [a,b]	0.092**	-0.060**	0.056*	-0.045	0.051	-0.073**	0.003	-0.037
	(0.037)	(0.029)	(0.063)	(0.028)	(0.041)	(0.037)	(0.030)	(0.035)
Labor force participation [a,b]	0.012***	0.003	0.000	-0.003	0.016***	0.002	0.018**	-0.002
	(0.004)	(0.004)	(0.008)	(0.007)	(0.005)	(0.005)	(0.008)	(0.007)
Military service dummy [a]	0.102	-0.189**	0.060	-0.217***	-0.025	-0.266**	-0.075	-0.082
	(0.089)	(0.087)	(0.090)	(0.083)	(0.142)	(0.121)	(0.115)	(0.102)
Polygamy dummy [a]	-0.943***	-0.421**	-0.677**	-0.618***	-1.828***	-0.569**	-0.428	-0.322
	(0.224)	(0.174)	(0.280)	(0.208)	(0.273)	(0.233)	(0.482)	(0.251)
GCC dummy [a]	0.623**	-0.332	0.685**	0.463	1.557***	-0.909***	0.162	-0.505
	(0.259)	(0.240)	(0.319)	(0.331)	(0.400)	(0.380)	(0.568)	(0.407)
Observations	10,143	10,143	10,143	10,143	7,892	7,892	7,892	7,892
Sq. Correlation coefficient	0.970	0.978	0.965	0.986	0.977	0.945	0.968	0.962

Notes. All regressions include fixed effects for origin countries and destination regions. Observations are weighted by the log of the bilateral migration stock. Superscript [a] denotes destination characteristics. Superscript [b] denotes that the variable is gender specific. OPSW = Bilateral stock of migrants provided by OPSW (2010). Superscripts ***,**,* denote statistical significance at 1, 5 and 10 percent, respectively. Robust standard errors are provided in parentheses.

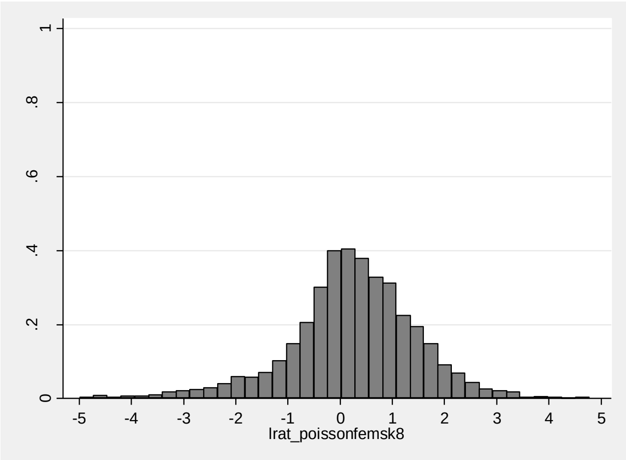
Table 3. In-sample comparison between actual and predicted stocks in 2000

	High-skilled females						High-skilled males					
	Obs	Corr	Migs	As %	Mean LR	Sd LR	Obs	Corr	Migs	As %	Mean LR	Sd LR
All	10,143	0.99	11.2	100.0	0.32	1.21	10,143	0.98	12.2	100.0	0.31	1.18
Less than 250	7,969	0.62	0.1	1.2	0.43	1.24	7,804	0.64	0.1	1.2	0.43	1.18
Above 250	1,514	0.98	10.8	96.4	0.05	0.63	1,646	0.98	11.8	96.7	0.08	0.66
Above 5000	287	0.98	8.9	79.5	-0.06	0.37	308	0.98	9.6	78.7	-0.05	0.40
Non-OECD>250	352	0.87	1.0	8.9	0.05	0.75	429	0.94	1.7	13.9	0.08	0.75
OECD>250	1,162	0.99	9.8	87.5	0.05	0.59	1,217	0.98	10.1	82.8	0.08	0.63
	Low-skilled females						Low-skilled males					
	Obs	Corr	Migs	As %	Mean LR	Sd LR	Obs	Corr	Migs	As %	Mean LR	Sd LR
All	10,143	0.99	26.4	100.0	0.24	1.28	10,143	0.99	27.8	100.0	0.18	1.25
Less than 250	7,241	0.56	0.1	0.5	0.30	1.22	7,303	0.58	0.1	0.5	0.24	1.21
Above 250	2,057	0.99	25.9	98.1	0.10	0.78	2,049	0.99	27.4	98.6	0.08	0.77
Above 5000	515	0.99	22.8	86.4	0.01	0.46	521	0.99	24.5	88.1	0.01	0.46
Non-OECD>250	582	0.99	7.0	26.5	0.05	0.81	625	0.99	9.8	35.3	0.04	0.76
OECD>250	1,475	0.99	18.9	71.6	0.13	0.77	1,424	0.99	17.6	63.3	0.10	0.77

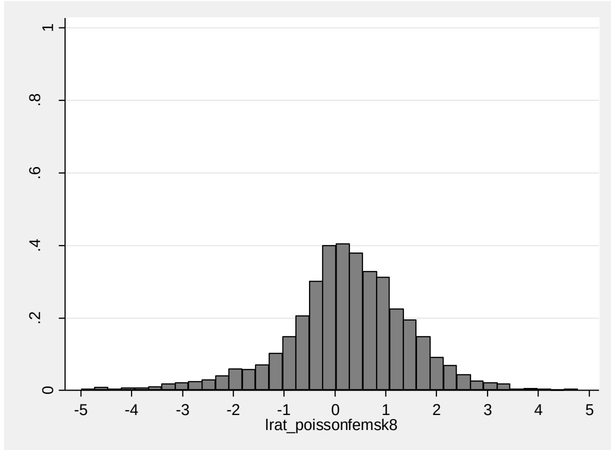
Notes: Column 'Obs' provides the number of observations; 'Corr' gives the correlation between our predictions and the actual migrant stocks; 'Migs' details the actual number of migrants (in millions); while 'As %' provides the percentage of the within sample migrants contained within each sample; 'Mean LR' gives the mean of the log of the ratio of actual-to-predicted migrant stocks; while the standard deviation of this logged ratio is provided in the 'Sd LR' column.

Figure 2. Distribution of the log-difference between actual and predicted stocks
High-skilled female migrants by corridor size in 2000

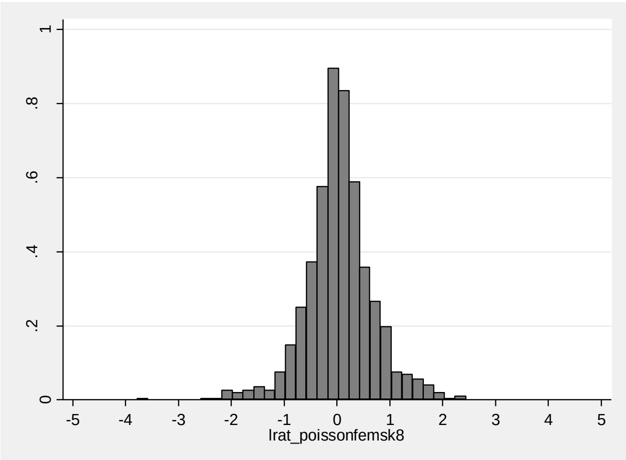
2.a. All corridors



2.b. Fewer than 250



2.c. Greater than 250



2.d. Greater than 5,000

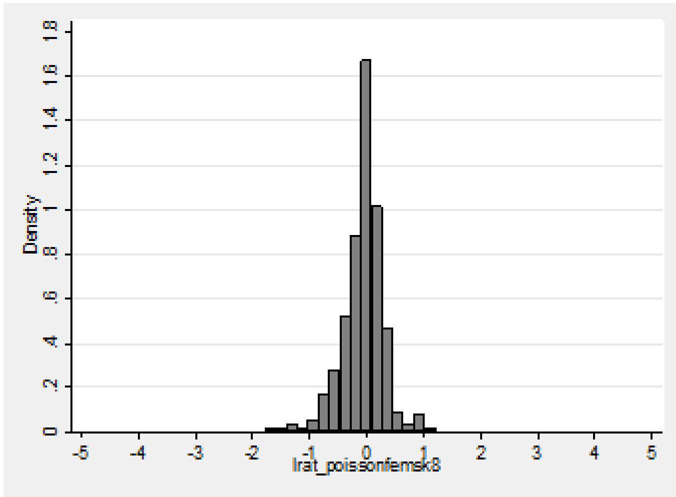


Table 4. Emigration patterns by country group, 1990 and 2000

	Total emigration			Emigration to OECD			Emigration to non-OECD		
	Stock (million)	College (%)	Women (%)	Stock (million)	College (%)	Women (%)	Stock (milion)	College (%)	Women (%)
Year 2000									
WORLD	100.5	26.1	48.8	59.3	35.3	51.0	41.2	12.7	45.5
OECD	32.3	30.0	50.4	29.1	31.0	50.8	3.2	21.1	46.3
HIGH	25.8	36.0	52.0	22.3	38.4	53.0	3.5	20.8	46.0
DEV	74.8	22.6	47.7	37.0	33.4	49.9	37.7	12.0	45.5
LOW	14.9	9.4	44.8	2.5	38.0	48.5	12.4	3.7	44.0
LDC	15.2	8.6	43.5	2.4	34.6	47.7	12.8	3.7	42.7
SIDS	4.5	34.3	55.1	4.0	37.1	54.9	0.5	10.1	56.8
USA	0.9	58.8	50.4	0.7	62.9	52.6	0.2	43.3	41.9
CANZ	1.5	57.1	54.0	1.4	57.6	54.3	0.1	47.3	47.1
EU27	20.2	31.5	51.9	17.7	33.1	52.4	2.5	19.6	48.4
GCC	0.4	22.9	35.4	0.0	65.2	39.7	0.4	17.1	34.8
LAC	15.6	25.0	50.2	14.0	26.4	50.1	1.5	11.9	51.4
SSA	14.1	8.3	45.6	2.2	43.1	47.5	11.9	1.8	45.2
CIS	10.6	24.1	55.8	2.4	42.1	58.2	8.2	18.8	55.1
INDIA	4.9	31.8	39.7	1.7	60.5	47.2	3.2	16.6	35.7
CHINA	4.1	28.1	52.4	1.7	46.7	53.0	2.4	15.0	51.9
MENA	8.4	23.8	38.5	4.2	29.9	43.0	4.2	17.6	33.9
Year 1990									
WORLD	80.2	20.4	47.4	42.6	29.5	50.7	37.7	10.0	43.6
OECD	25.9	25.9	50.7	23.3	26.9	51.6	2.6	17.3	42.9
HIGH	23.6	28.6	51.8	20.5	30.5	52.8	3.2	15.9	45.1
DEV	56.6	17.0	45.5	22.1	28.6	48.7	34.5	9.5	43.5
LOW	12.3	6.9	44.3	1.4	33.7	45.6	10.9	3.4	44.1
LDC	12.8	6.4	43.4	1.4	30.2	45.1	11.4	3.5	43.2
SIDS	2.9	32.3	53.4	2.6	34.6	53.6	0.2	7.6	51.7
USA	0.8	50.6	49.9	0.6	53.8	53.0	0.2	39.6	39.0
CANZ	1.3	46.0	55.8	1.2	46.4	56.3	0.1	37.0	46.3
EU27	19.0	24.8	51.5	16.9	26.0	52.2	2.1	15.3	45.8
GCC	0.3	19.0	31.1	0.0	64.8	35.6	0.3	14.9	30.7
LAC	8.2	24.7	50.1	7.0	27.4	50.8	1.2	9.0	46.5
SSA	10.9	5.8	46.2	1.2	39.6	44.3	9.7	1.6	46.5
CIS	10.0	14.9	52.7	1.8	20.8	56.3	8.2	13.6	51.9
INDIA	4.8	19.3	37.5	1.0	45.5	47.0	3.8	12.6	35.1
CHINA	3.2	17.8	57.5	0.9	40.0	50.2	2.3	9.2	60.4
MENA	6.8	19.2	35.0	3.2	23.8	41.5	3.6	15.1	29.4

Notes. Colomn ‘Stock’ gives the aggregate stock of emigrants in millions ; ‘College’ gives the percentage of high-skilled emigrants ; ‘Women’ gives the percentage of female emigrants.. For high-income (HIGH), developing (DEV) and low-income countries (LOW), we use the World Bank classification. Least developed countries (LDC) and small island developing states (SIDS) are defined by the United Nations. EU27: 27 countries of the European Union, USA: United States of America, CANZ: Canada + Australia + New Zealand; CIS: Commonwealth of independent States of the former USSR, MENA: Middle East and Northern Africa, SSA: Sub-Saharan Africa. Each country only belongs to one geographical group.

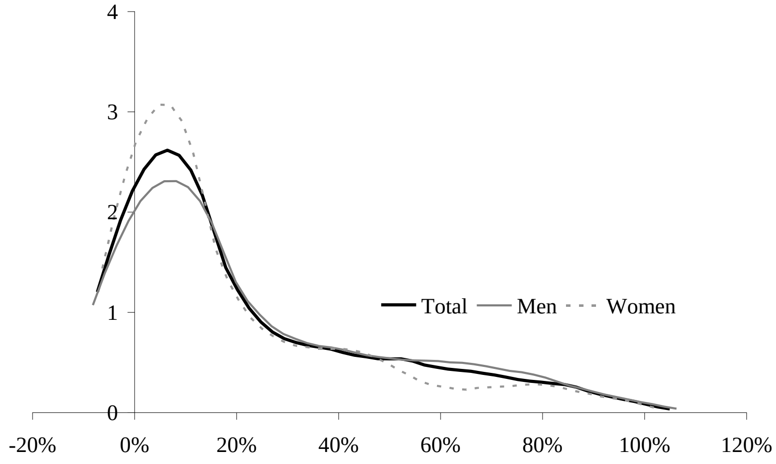
Table 5. High-skilled emigration rates, 1990 and 2000

	Gross high-skilled emigration rate			Net high-skilled emigration rates		
	To all	To OECD	To non-OECD	Total	Men	Women
Year 2000						
WORLD	7.3	5.9	1.5	0.0	0.0	0.0
OECD	4.8	4.5	0.3	-5.6	-5.4	-5.7
HIGH	4.7	4.4	0.4	-6.7	-6.9	-6.5
DEV	10.5	7.7	2.8	8.2	7.1	9.7
LOW	19.4	13.1	6.3	15.1	13.1	20.0
LDC	20.2	12.9	7.3	16.6	14.8	21.0
SIDS	40.9	39.7	1.2	34.9	29.2	41.1
USA	0.6	0.5	0.1	-11.6	-12.1	-11.2
CANZ	7.2	6.9	0.3	-30.9	-32.4	-29.4
EU27	9.5	8.8	0.7	2.2	2.3	2.1
GCC	10.9	3.7	7.1	-113.1	-260.4	-34.0
LAC	12.1	11.6	0.6	10.9	9.9	12.0
SSA	15.0	12.2	2.7	10.5	8.7	14.3
CIS	8.1	3.2	4.9	3.3	2.7	4.1
INDIA	6.5	4.3	2.2	5.8	5.2	7.3
CHINA	5.4	3.7	1.7	5.3	3.8	9.3
MENA	17.6	11.1	6.5	9.3	10.5	6.9
Year 1990						
WORLD	6.8	5.2	1.6	0.0	0.0	0.0
OECD	4.6	4.3	0.3	-4.0	-3.8	-4.3
HIGH	4.8	4.4	0.4	-4.8	-4.7	-4.8
DEV	9.7	6.4	3.3	6.8	6.0	8.3
LOW	22.0	12.3	9.6	15.0	13.9	18.1
LDC	23.9	12.4	11.5	15.5	14.1	19.8
SIDS	41.9	41.1	0.8	38.1	33.3	43.8
USA	0.7	0.5	0.1	-10.0	-9.4	-10.9
CANZ	6.6	6.3	0.2	-28.9	-30.5	-27.0
EU27	9.1	8.5	0.6	3.8	3.8	3.9
GCC	11.4	3.2	8.2	-107.7	-193.1	-38.9
LAC	10.7	10.1	0.6	9.5	8.5	10.7
SSA	17.1	13.1	4.1	10.9	9.2	15.5
CIS	6.3	1.6	4.7	2.1	1.9	2.4
INDIA	5.8	2.8	3.0	4.4	4.0	5.4
CHINA	4.7	3.0	1.8	4.6	3.0	11.6
MENA	22.5	13.0	9.5	12.5	12.4	12.8

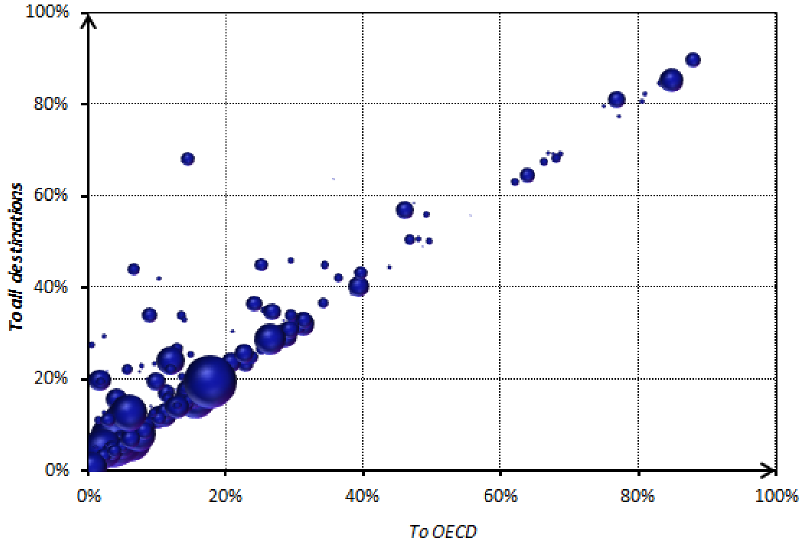
Notes. Colomn 'Stock' gives the aggregate stock of emigrants in millions ; 'College' gives the percentage of high-skilled emigrants ; 'Women' gives the percentage of female emigrants.. For high-income (HIGH), developing (DEV) and low-income countries (LOW), we use the World Bank classification. Least developed countries (LDC) and small island developing states (SIDS) are defined by the United Nations. EU27: 27 countries of the European Union, USA: United States of America, CANZ: Canada + Australia + New Zealand; CIS: Commonwealth of independent States of the former USSR, MENA: Middle East and Northern Africa, SSA: Sub-Saharan Africa. Each country only belongs to one geographical group.

Figure 3. Distribution of high-skilled emigration rates

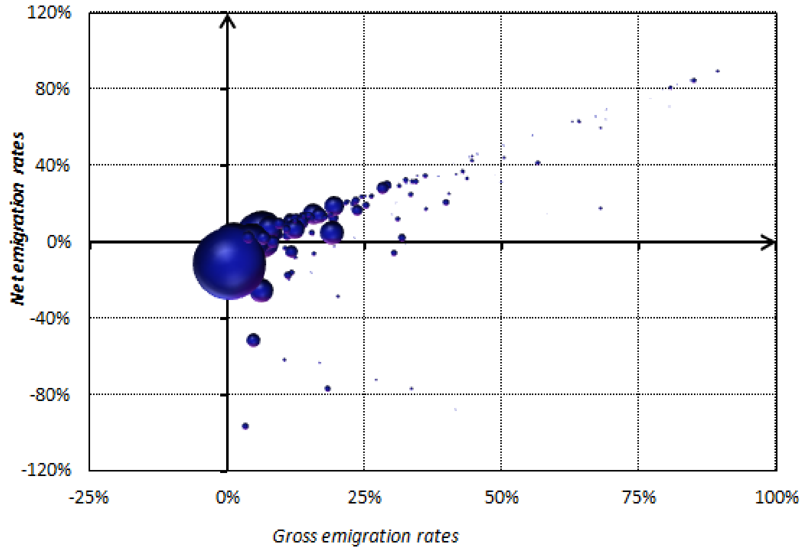
3.a. Density of “non-OECD to total” ratio of emigration rates



3.b. High-skilled emigration rates to OECD and to all destinations in 2000



3.c. Net versus gross emigration rates in 2000



Notes. On Figures 3.b and 3.c, each country is a represented by a bubble, the size of which is proportional to the high-skilled emigration stock.

.

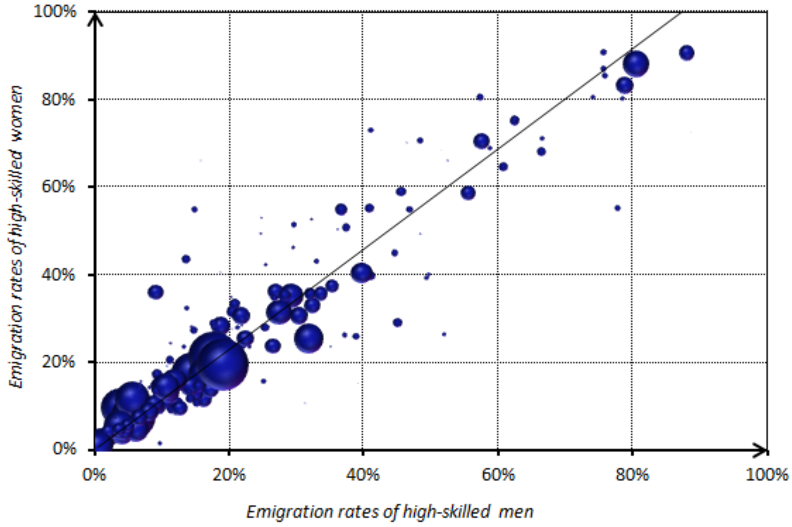
Table 6. Brain drain: most and least affected countries

Largest high-skilled net emigration rates							Lowest high-skilled net emigration rates						
	2000			1990				2000			1990		
Country	Net (%)	Gross (%)	non-OECD (%)	Net (%)	Gross (%)	non-OECD (%)	***Country***	Net (%)	Gross (%)	non-OECD (%)	Net (%)	Gross (%)	non-OECD (%)
Jamaica	84.7	85.0	0.2	85.6	85.6	0.5	United Arab Emirates	-325.8	5.3	48.3	-105.5	2.0	66.0
Haiti	80.7	80.9	4.9	72.5	72.5	3.0	Saudi Arabia	-96.5	3.3	45.0	-134.5	4.8	54.6
Liberia	55.8	55.8	11.7	58.8	61.0	18.5	Israel	-77.1	18.4	18.0	-18.2	14.1	20.9
Sierra Leone	50.4	50.4	4.5	48.6	48.6	17.2	Kuwait	-77.1	33.8	59.7	-45.8	35.0	72.4
Eritrea	45.7	45.7	35.4	48.3	48.3	47.3	Oman	-72.6	27.3	97.7	-57.9	27.8	98.3
Somalia	44.8	44.8	23.0	34.4	34.4	33.7	Libya	-62.2	10.5	17.0	-191.7	28.4	32.9
Laos	44.0	50.4	7.1	34.4	47.0	3.5	Australia	-51.7	4.9	12.1	-47.5	3.6	13.5
Afghanistan	42.4	44.7	43.5	26.7	31.9	66.2	Canada	-25.8	6.3	1.9	-23.0	6.4	1.3
Lebanon	41.5	56.6	18.7	53.1	67.3	20.8	Switzerland	-17.6	11.3	8.6	-12.6	8.1	8.8
Kenya	36.7	43.1	7.9	47.9	49.8	8.6	Singapore	-16.1	11.9	19.0	0.6	10.8	10.2
Uganda	35.2	41.9	12.9	42.3	43.8	19.2	United States	-11.6	0.6	15.3	-10.0	0.7	17.5
Bosnia Herzegovina	34.7	36.2	33.1	30.2	33.5	28.8	Sweden	-6.7	5.2	3.1	-4.0	4.2	1.5
Congo, Rep. of the	34.4	34.7	25.9	11.8	20.2	16.8	Latvia	-6.5	16.1	26.3	-54.2	16.6	32.6
Yemen	33.0	43.8	84.6	32.9	93.7	74.1	Cote d'Ivoire	-6.4	11.0	39.9	-14.6	9.0	38.4
El Salvador	32.3	32.7	3.6	33.1	35.3	9.4	New Zealand	-5.7	30.5	3.4	-19.6	25.4	4.2
Nicaragua	31.7	33.8	12.7	28.8	32.5	15.5	Netherlands	-5.1	11.9	5.4	-2.7	11.9	4.5
Sri Lanka	31.5	34.5	22.1	32.6	38.4	35.0	Paraguay	-4.6	5.9	29.9	0.2	4.2	5.4
Togo	30.7	32.8	57.2	27.0	35.7	67.9	Norway	-2.9	7.0	2.0	0.8	7.8	1.4
Cuba	29.3	29.3	2.6	31.8	31.8	2.9	Belgium	-2.8	6.5	6.6	-0.7	5.4	6.5
Macedonia	29.0	31.4	4.7	26.5	30.6	7.9	Papua New Guinea	-2.6	19.6	1.7	-50.3	34.5	0.1
Vietnam	28.0	28.5	6.6	24.5	24.9	5.3	Spain	-2.5	4.3	14.2	-0.8	3.4	12.0
Jordan	24.9	33.7	73.0	35.4	53.0	78.6	France	-2.4	4.0	14.7	-0.8	3.5	20.9
Armenia	24.1	26.5	51.0	21.3	22.2	53.3	Germany	-0.1	6.6	5.5	2.4	6.8	4.6
Zambia	23.9	25.3	40.8	29.2	31.8	45.4	Russia	0.0	4.5	52.3	0.3	4.0	71.4
Honduras	23.5	26.3	4.1	20.8	27.9	18.2	Costa Rica	0.0	8.1	6.2	-5.3	11.1	13.8

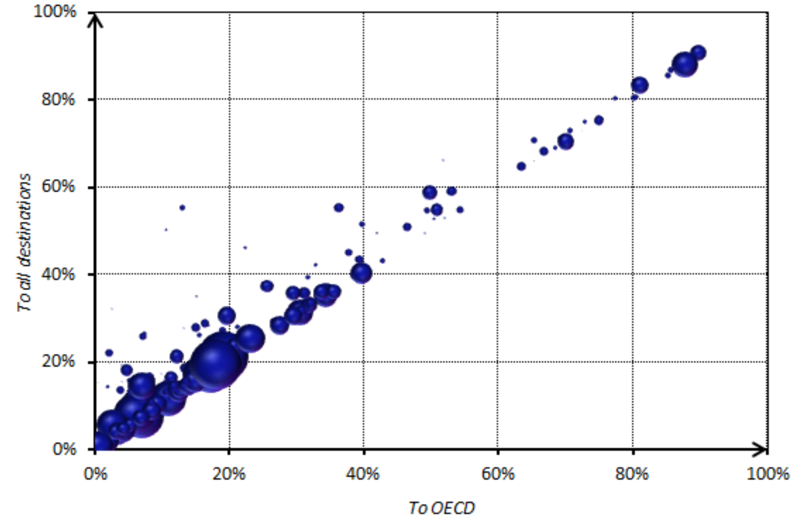
Notes. Only countries with labor force above one million are included. Countries' ranking is based on net emigration rates in 2000. The non-OECD share measures the share of non-OECD countries in gross emigration of college graduates.

Figure 4. Distribution of emigration rates of high-skilled women

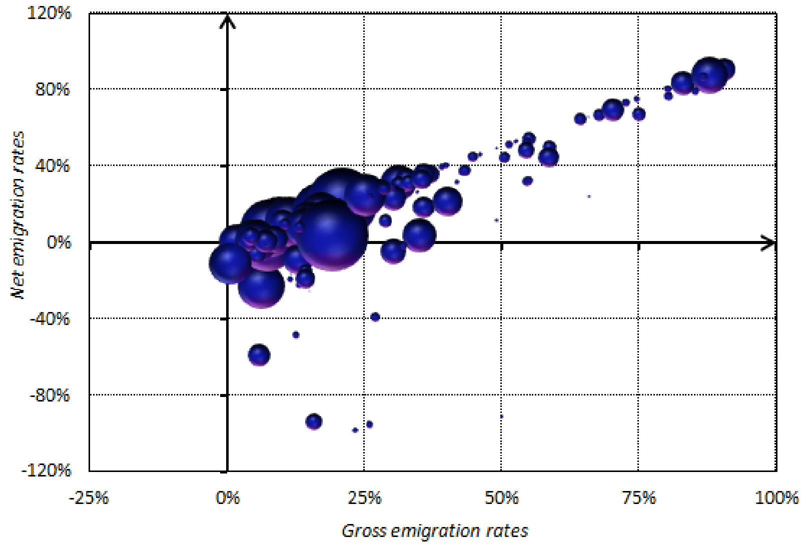
4.a. Emigration rates of high-skilled women and men



4.b. Emigration rates of high-skilled women to OECD and to all destinations in 2000



4.b. Net versus gross emigration rates of high-skilled women in 2000



Notes. On Figures 2.b and 2.c, each country is a represented by a bubble, the size of which is proportional to the emigration stock of high-skilled women in 2000.

.

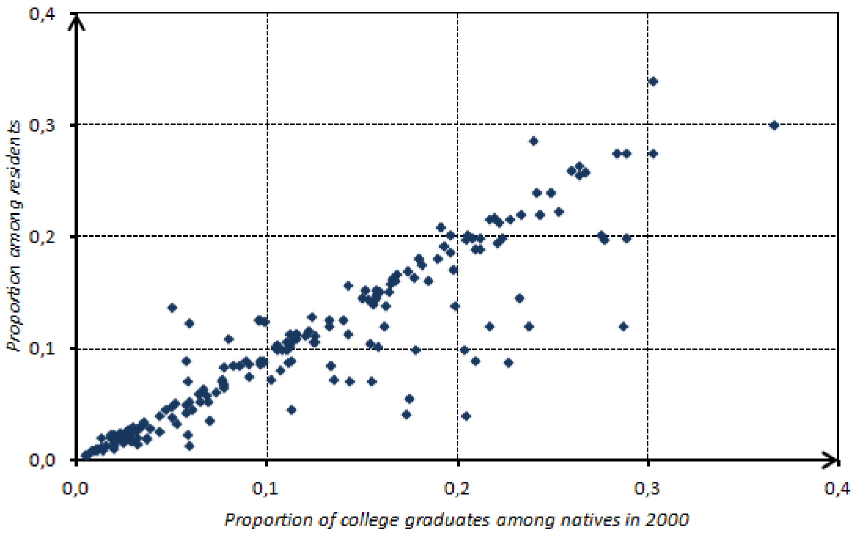
Table 7. Women's brain drain: most and least affected countries

Highest net high-skilled emigration rates							Lowest net high-skilled emigration rates						
	2000			1990				2000			1990		
Country	Net (%)	Gross (%)	non-OECD (%)	Net (%)	Gross (%)	non-OECD (%)	***Country***	Net (%)	Gross (%)	non-OECD (%)	Net (%)	Gross (%)	non-OECD (%)
Jamaica	87.7	88.0	0.2	87.6	87.6	0.4	United Arab Emirates	-213.6	3.3	36.6	-125.3	1.5	51.9
Haiti	83.2	83.2	2.6	78.0	78.0	1.0	Kuwait	-95.5	26.1	39.9	-93.6	18.2	40.9
Sierra Leone	72.8	72.8	2.8	73.7	73.7	17.8	Israel	-94.4	16.0	11.6	-21.5	10.5	5.4
Liberia	70.5	70.5	7.2	68.9	68.9	19.0	Australia	-59.6	6.0	10.8	-66.2	5.5	12.6
Afghanistan	54.0	55.1	34.0	44.7	44.7	48.1	Libya	-48.7	12.7	17.0	-54.8	22.8	30.8
Cameroon	52.4	54.7	9.6	30.6	30.6	16.9	Oman	-26.1	15.1	96.5	-22.5	12.9	97.3
Congo, Rep.	51.3	51.3	22.3	26.8	34.2	14.3	Canada	-23.1	6.4	1.5	-19.0	6.7	0.9
Laos	49.7	58.9	9.6	35.5	53.5	6.2	Nepal	-22.6	13.2	60.5	-185.0	46.6	64.1
Kenya	48.1	54.7	6.9	58.2	59.9	4.7	Switzerland	-19.1	14.4	6.4	-11.0	10.9	5.5
Togo	46.1	46.1	51.4	48.6	48.6	75.7	Singapore	-15.1	14.3	16.6	5.1	13.8	7.0
Somalia	44.9	44.9	15.7	33.1	33.1	16.5	Saudi Arabia	-13.2	1.4	41.3	-19.1	1.6	41.8
Lebanon	44.3	58.6	14.9	57.9	68.9	13.6	United States	-11.2	0.6	12.1	-10.9	0.7	12.2
Uganda	44.3	50.7	8.1	56.6	58.1	9.4	Cote d'Ivoire	-11.0	8.3	23.8	-22.9	4.0	1.3
Eritrea	39.1	39.1	18.7	42.9	42.9	22.5	Netherlands	-9.0	13.0	4.0	-7.0	13.8	2.8
Congo, Dem.	37.4	43.4	9.2	43.6	60.3	22.4	Latvia	-8.0	19.1	25.5	-56.9	19.0	34.0
Bosnia and Herzegovina	35.8	37.2	31.0	31.4	33.4	24.0	Sweden	-6.2	5.5	2.5	-3.6	4.7	0.7
Nigeria	35.3	35.9	6.1	15.6	15.6	13.8	New Zealand	-4.8	30.5	3.0	-14.2	24.9	3.6
Nicaragua	33.2	35.6	12.1	31.7	34.1	12.3	Norway	-3.1	7.0	1.3	0.6	8.1	0.8
Sri Lanka	32.8	35.6	16.7	29.1	31.6	16.7	Paraguay	-2.8	6.8	33.5	2.4	4.4	3.5
Macedonia	32.7	35.3	4.4	31.2	35.2	6.1	Spain	-2.7	4.1	10.8	-1.5	3.3	7.9
El Salvador	32.6	32.9	2.8	33.9	35.9	6.5	Belgium	-2.4	6.9	4.1	-0.8	5.9	4.0
West Bank Gaza	32.3	55.2	76.2	38.3	45.0	76.9	France	-1.8	3.9	11.6	0.0	3.5	13.3
Rwanda	31.4	42.1	21.9	-40.4	75.7	51.8	Turkmenistan	-1.2	7.2	77.0	6.7	7.7	87.8
Cuba	31.3	31.3	2.1	32.8	32.8	2.3	Russia	0.3	5.4	50.5	0.3	4.6	71.2
Honduras	30.8	33.3	3.2	33.1	37.3	13.6	Japan	0.4	1.9	7.2	0.9	1.8	3.5

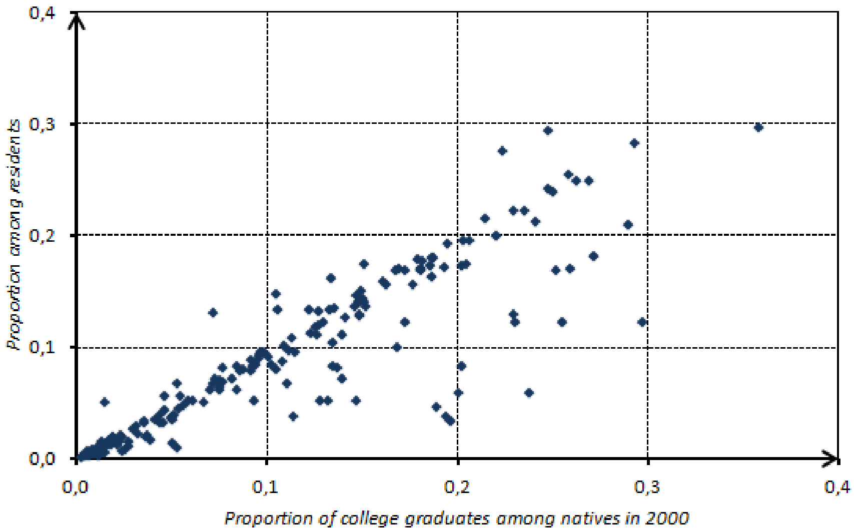
Notes. Only countries with labor force above one million are included. Countries' ranking is based on net emigration rates in 2000. The non-OECD share measures the share of non-OECD countries in gross emigration of college graduates.

Figure 5. Human capital among natives and residents in 2000

5.a. Women and men together



5.b. Women only



Note: Human capital is measured by the proportion of college graduates in the population aged 25 and more.

Table A1. Migration data for additional destinations

Country	Source	1990		2000	
		Total	High-skilled	Total	High-skilled
Saudi Arabia	Labor Force Survey[c]	2,864,310	401,003	3,101,890	582,253
Israel	Central Bureau of Statistics	1,530,890	228,630	1,512,645	512,785
Hong Kong	Census and Statistics Dep.	-	-	1,883,552	294,419
Un Arab Emirates	Labor Force Survey[c]	790,516	148,387	1,160,699	217,874
Philippines	IPUMS International[b]	77,077	28,573	635,696	194,286
South Africa	Statistics South Africa	635,114	101,877	795,069	174,876
Singapore	Statistics Singapore	397,365	30,210	512,726	137,816
Kuwait	Labor Force Survey[c]	668,926	125,563	668,926	125,563
Dominican Rep	United Nations CEPAL[a]	-	-	217,334	94,613
Oman	Labor Force Survey[c]	307,010	57,629	411,692	77,278
Estonia	Statistics Estonia	407,407	114,283	233,166	72,609
Latvia	Latvia Statistics	689,314	110,324	415,067	69,919
Croatia	Central Bureau of Statistics	388,596	46,558	498,918	68,891
Brazil	IPUMS International[b]	346,699	67,599	302,367	67,817
Argentina	IPUMS International[b]	742,467	92,715	694,919	62,686
Kyrgyztan	IPUMS International[b]	-	-	314,940	46,880
Qatar	Labor Force Survey[c]	198,450	37,251	247,230	46,407
Malaysia	IPUMS International[b]	-	-	1,006,800	45,900
Lithuania	Statistics Lithuania	272,779	41,498	204,097	42,573
Venezuela	IPUMS International[b]	520,170	19,116	515,612	39,066
Ivory Coast	Institut National de Stat.	3,294,870	30,147	3,942,022	36,068
Romania	IPUMS International[b]	92,801	29,435	87,233	32,552
Costa Rica	I.N. Estadistica y Censos	235,750	29,976	175,527	29,312
Bahrain	Labor Force Survey[c]	153,576	28,828	153,576	28,828
Bulgaria	National Statistical Institute	16,411	4,776	77,056	26,383
Iraq	IPUMS International[b]	-	-	116,620	21,818
Paraguay	United Nations CEPAL[a]	-	-	131,397	20,488
Slovenia	Statistical Office Slovenia	166,187	20,296	153,827	18,220
Morocco	Haut Commissariat au Plan	-	-	40,023	17,684
Cyprus	Cyprus Statistics	23,679	8,787	43,263	17,332
Kenya	IPUMS International[b]	150,800	13,160	206,580	16,900
Bolivia	IPUMS International[b]	-	-	48,220	15,780
Panama	IPUMS International[b]	-	-	63,830	13,050
Uganda	IPUMS International[b]	274,905	922	199,050	11,910
Macedonia	State Statistical Office	81,106	10,614	77,567	11,826
Belarus	IPUMS International[b]	50,931	10,392	54,660	11,070
Malta	National Statistics Office	12,613	5,279	19,009	8,524
Chile	IPUMS International[b]	44,590	3,080	74,430	6,490
Colombia	IPUMS International[b]	45,100	3,400	52,793	5,891
Honduras	United Nations CEPAL[a]	-	-	18,042	5,826
Guinea	IPUMS International[b]	-	-	129,490	5,600
Nicaragua	United Nations CEPAL[a]	-	-	42,163	4,936
Rwanda	IPUMS International[b]	101,652	9,296	134,670	4,900
Trinidad and Tobago	United Nations CEPAL[a]	-	-	31,897	2,699
Mongolia	IPUMS International[b]	-	-	5,480	1,940
Belize	United Nations CEPAL[a]	-	-	21,954	1,224

Notes. [a] United Nations' Economic Commission for Latin America and the Caribbean (http://www.cepal.org). [b] See Minnesota Population Center (2010) and https://international.ipums.org. [c] Data for GCC countries: for Saudi Arabia, see *Population and Social Statistics* at http://www.cdsi.gov.sa; for the United Arab Emirates, see *Statistic Reports-Census 2005* at http://www.economy.ae; for Qatar, see *Labour Force Sample Survey* at http://www.qsa.gov.qa; for Bahrain, see *Labour Market Indicators* at http://blmi.lmra.bh; for Oman, see Periodic *Labour Force Survey* at http://www.moneoman.gov.om; and for Kuwait, see *Microdata of the Labor Force Survey* at http://scs.mop.gov.kw.

Table A2. Description of Explanatory Variables

Variable	Source	Description
Common border	CEPII[a]	Dummy equal to 1 if a country pair share a land border
Distance	CEPII[a]	Measure of geodesic distance between country pair's main cities
Common language	CEPII[a]	Dummy equal to 1 if a country pair shares a common official language
Former colony	CEPII[a]	Dummy equal to 1 if a country pair share a colonial history
OPSW bilateral stock	OPSW (2010)	Total migrant stock recorded between origin i and destination j
Some English	CIA World Factbook[b]	Dummy equal to 1 if a destination country speaks some English
GDP per capita	Penn World Tables[c]	Per capita income of the destination country in PPP
Total fertility	World Development Indicators	Total fertility rate (in log) in the destination country
Skill destination workforce	DLM (2009)	Share of the destination country workforce that are tertiary educated (by gender)
Total labor force	DLM (2009)	Population aged 25 and over in the destination country (by gender)
Labor force participation	World Development Indicators	Labor force participation rate in the destination country (by gender)
Military service dummy	Own calculation	Dummy equal to 1 if military service is compulsory in the destination country
Polygamy dummy	Own calculation	Dummy equal to 1 if polygamy is legally or socially accepted in the destination country
GCC dummy	Own calculation	Dummy equal to 1 if a destination country belongs to GCC

Notes: a See: http://www.cepii.fr/anglaisgraph/bdd/distances.htm, see Clair et al. (2004).
[b] See: https://www.cia.gov/library/publications/the-world-factbook.
[c] See: http://pwt.econ.upenn.edu.

Institut de Recherches Économiques et Sociales
Université catholique de Louvain

Place Montesquieu, 3
1348 Louvain-la-Neuve, Belgique

ISSN 1379-244X D/2012/3082/022