www.elsevier.com/locate/worlddev

World Development Vol. 65, pp. 6–26, 2015
0305-750X/© 2014 Published by Elsevier Ltd.

http://dx.doi.org/10.1016/j.worlddev.2014.04.004

A Global Assessment of Human Capital Mobility: The Role of Non-OECD Destinations

ERHAN ARTUC [a], FRÉDÉRIC DOCQUIER [b,c], ÇAGLAR ÖZDEN [a,d] and CHRISTOPHER PARSONS [e,*]

[a] *World Bank, Washington, DC, USA*
[b] *IRES Université Catholique de Louvain, Belgium*
[c] *FNRS, Belgium*
[d] *IZA, Bonn, Germany*
[e] *University of Oxford, UK*

Summary. — Discussions of high-skilled mobility typically evoke migration patterns from poorer to wealthier countries, which ignore movements to and between developing countries. This paper presents, for the first time, a global overview of human capital mobility through bilateral migration stocks by gender and education in 1990 and 2000, and calculation of nuanced brain drain indicators. Building on newly collated data, we use a novel estimation procedure based on a pseudo-gravity model. We identify key determinants of international migration, which we subsequently use to impute missing data. Non-OECD destinations account for one-third of skilled-migration, while OECD destinations are declining in relative importance.

Key words — international migration, labor mobility, brain drain

1. INTRODUCTION

Among the various dimensions of international migration, movements of the highly skilled are arguably the most topical. On the one hand, governments of more developed countries are implementing policies to attract the best and the brightest in an increasingly competitive market for skills. On the other hand, many poorer countries, especially those already suffering from low levels of human capital, are deeply concerned about retaining their most skilled workers, whose absence ultimately impinges upon their long-term economic and political development. Until now, the literature has almost exclusively examined high-skilled movements to OECD countries, often termed the "brain drain." Even a casual observation of basic migration patterns, however, indicates that such a focus fails to capture the complete global picture.

The absence of detailed and high quality data is the main obstacle that prevents us from properly quantifying the extent of skill mobility across the world. These data shortcomings not only impede many important avenues of research, but in light of the paucity of immigration and emigration flow data by skill level, also militate against countries' ability to assess their net human capital situation and thus the effectiveness of their immigration, education, and labor market policies. [1] This paper is the first to seriously address this issue, by first developing a global overview of human capital mobility and then subsequently by introducing refined brain drain indicators, which, in comparison with the existing literature, provide superior estimates of gross and net human capital levels across the world.

There have been several efforts to analyze bilateral migration patterns. The Eurostat database [2] provides data on the size of migration flows, by age, gender, and country of citizenship, but solely between EU member states and numerous missing observations exist. More broadly, Özden, Parsons, Schiff, and Walmsley (2011) referred to as OPSW henceforth, construct five 226×226 comprehensive matrices of origin–destination stocks that correspond to the last five completed census rounds, thereby extending the work of Parsons, Skeldon, Walmsley, and Winters (2007). However, while OPSW significantly broadens the time, gender, and geographical coverage of the available data, different skills or education levels are not distinguished.

Another set of studies investigates the education structure of migration, but only for a limited set of destination countries for which data are more readily available. Docquier and Marfouk (2004, 2006) and Dumont and Lemaître (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 analyses.

Existing data sets of bilateral migrant stocks disaggregated by education level only capture the size and structure of migration to OECD destinations. This is an important limitation, since migration to non-OECD countries is significant. Figure 1 shows that the share of non-OECD destination countries in the world immigration stock has gradually decreased since the 1960s (from 57% to 49%). Nevertheless, non-OECD

* We would like to extend our thanks to the Knowledge for Change Program of the World Bank and 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. Final revision accepted: April 11, 2014.

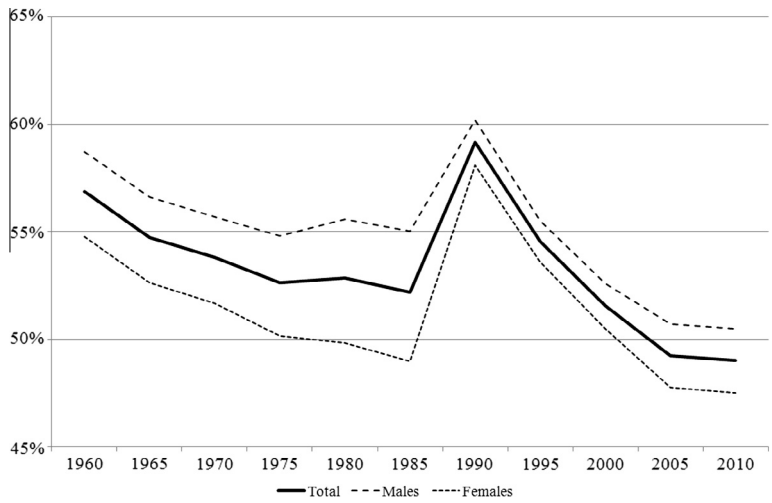


Figure 1. *Share of non-OECD destinations in the world migration stock (Data by gender, 1960–2010) Source: United Nations Population Division (2007, 2012).*

countries still host about half of all current international migrants. This share is not homogenous across gender since it is larger for men (51.3% in 2000) than for women (48.7%). Countries such as Russia, Ukraine, India, and Pakistan attract large numbers of migrants, mostly from neighboring countries and as a result of political events that changed national boundaries. 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 SAR) are among the most important non-OECD destinations. Omitting these destinations from any analysis results in an important piece of the global puzzle remaining missing, thereby limiting our understanding of the full nature of international human capital mobility.

In this paper, we perform, for the first time, a global analysis of bilateral migration patterns by gender and for two education levels, i.e., for four labor types. Compared to previous analyses, we account for migration to all non-OECD country destinations by introducing new data and utilizing appropriate estimation methods where actual bilateral data are missing. Furthermore, we are able to refine existing measures of immigration and emigration rates by expressing immigrant and emigrant stocks relative to a more appropriate measure of the labor force, the *natural* labor force, i.e., the number of workers from a particular origin country regardless of where they currently reside.

Our analysis shows that migration to non-OECD countries increased at a slower pace (+23%) than migration to the OECD (+39%) during 1990–2000. Nevertheless, these former groups constitute about 47% of the world adult migration stock, which is characterized by both lower shares of college graduates (approximately half the level of migration to OECD countries) and women. The selection on skills is particularly pronounced in the case of least developed countries, increasing with regional income levels and for most global regions during 1990–2000. These patterns demonstrate the continued and increasing attractiveness of OECD destinations for high-skilled and female 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 favored by female migrants, the extent of this selection on gender decreased during 1990–2000, which highlights the rising appeal of non-OECD destinations for female migrants. Emigration to non-OECD countries accounts for about one-third of the total brain drain from low-income and the least developed countries and adding non-OECD destinations increases the high-skilled emigration rate of 32 countries by more than 50%. These countries 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 Persian Gulf countries. The influence of our introducing additional countries on female high-skilled emigration, however, is less pronounced, given the continued tendency for female migrants to migrate to OECD countries.

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 high-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 forces are typically more highly educated than the workforce that resides in those countries. In other words, high-skilled immigration for such countries, fails to compensate for the skill losses endured when college-educated natives move abroad.

Before delving into the details of the empirical exercise and our analysis of the 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%) have college education, and 30.2 million (51%) are women. For 1990, we identify 42.5 million migrants to OECD countries of which 30% are highly educated and 51% are women. The third and fourth columns show the data obtained or estimated for non-OECD countries. There are 52.6 million migrants, of which 7.9 million (15%) are highly educated and 24.3 million (46%) are females in 2000. For 1990, we identify 42.7 million migrants, including 8.7% highly educated and 45% 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

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

	Total (million)	To OECD (million) [a]	To non-OECD[a]		Including imputed stocks	
			(million)	(%)[b]	(million)	(%)[b]
Year 2000						
Total	111.9	59.3	52.6	47.0	16.7	14.9
College graduates	28.8	20.9	7.9	27.4	2.5	8.7
Less educated	83.1	38.3	44.7	53.9	14.2	17.1
Males	57.4	29.0	28.4	49.4	8.7	15.1
College graduates	15.1	10.6	4.5	30.0	1.4	9.0
Less educated	42.3	18.4	23.8	56.4	7.3	17.3
Females	54.5	30.2	24.3	44.5	8.0	14.7
College graduates	13.7	10.3	3.3	24.4	1.1	8.3
Less educated	40.8	19.9	20.9	51.3	6.9	16.8
Year 1990						
Total	85.3	42.5	42.7	50.1	30.5	35.7
College graduates	16.3	12.6	3.7	22.8	2.4	14.9
Less educated	69.0	30.0	39.0	56.6	28.1	40.7
Males	44.4	21.0	23.4	52.7	15.6	35.2
College graduates	9.0	6.7	2.3	25.2	1.4	15.1
Less educated	35.4	14.3	21.2	59.7	14.3	40.3
Females	40.9	21.6	19.3	47.3	14.8	36.3
College graduates	7.3	5.9	1.4	19.8	1.1	14.6
Less educated	33.6	15.7	17.9	53.3	13.8	41.0

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

migrants imputed across unobserved corridors. For 2000, imputed values represent 16.7 million migrants in the 90 destination countries for which actual data are not available. Although imputed values account for 15% of the total migration stock in 2000, the share is around 8.7% for college-educated migrants. In other words, over 90% of college-educated migrants are captured by our raw data and we believe that our imputation strategy should therefore not adversely effect our overall measurement of high-skilled emigration, thereby diluting our conclusions.

Overall, the resulting migration matrices identify 111.9 million migrants (age 25+) in 2000 which represents about 63% of the 177.4 million migrants (age 0+) recorded in the United Nations database, and 70% of the 160.1 million migrants (again age 0+) recorded in OPSW for those 190 countries that appear in our matrices.[3] 28.8 million of this migrant stock have college education and 54.5 million are women. For 1990, we identify 85.3 million migrants (aged 25+), including 16.3 million high-skilled migrants and 40.9 million women. Our data show that the overall migrant stock increased by 31% during 1990–2000, while the stock of high-skilled migrants increased by 77%. As a result, the share of high-skilled in the overall migrant stock increased from 19% to 26%. The share of women increased from 48% to 49%, a result in part driven by the increased feminization of migration to non-OECD countries.

The remainder of this paper is organized as follows. Section 2 describes our data collection, while our econometric strategy and the accompanying results are presented in Section 3. In the following section, we introduce our nuanced brain drain indicators before we present our global assessment of human capital mobility in Section 5. Finally, we conclude.

2. DATA COMPILATION

The first contribution of the paper is in compiling a more complete global data set of bilateral migrant stocks, disaggregated by education level and gender for the years 1990 and 2000, by including as many developing destination countries in our sample for which data are available. Our data collection builds upon the previous database of DLM, which documents migrant stocks disaggregated by education levels to 30 OECD destination countries. Our methodology in this section is a direct extension of this earlier work. We add four new OECD members (for both 1990 and 2000), 66 non-OECD destinations in 2000 and 27 non-OECD countries in 1990 for which comparable data could be found. 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 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.

Subsequently, we use the primary data from these 190×100 and 190×61 matrices, for 2000 and 1990 respectively, to make out-of-sample predictions for those destination countries for which data are missing. Taken together, the raw and imputed data comprise 190 countries in 1990 and 2000 (denoted by $j = 1, \ldots, J$) 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 labor migration. The full data set in turn facilitates, for the first time, a global analysis of human capital mobility over time using nuanced and improved brain drain indicators as described in Section 4.

(a) *Migration data for OECD countries*

Our starting point when constructing our matrices is the Docquier, Lowell, Marfouk (DLM) data set, which 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 data set omitted data for member states that subsequently joined the OECD in 2010 however

(namely Chile, Estonia, Israel, and Slovenia) and so in this paper, we augment the original OECD data from the DLM data set 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. Data sources for these four destination countries are presented in Table 7.

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

(i) One hundred and ninety origin countries in both 1990 and 2000 are distinguished. Starting with the 192 UN member states, we aggregate the Republic of 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 Taiwan (China), Hong Kong SAR, Macao, and the Palestinian Territories are added as individual entries to the country list. We drop five countries (Nauru, Palau, Tuvalu, Belize, and the Holy See) due to their small size and their absence in the data of some destination countries.

(ii) Migration is predominantly measured on the basis of country of birth as opposed to citizenship, since our goal is to have a consistent definition over time. Whereas individuals' country of birth is predominantly 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.

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

(iv) Along with the gender dimension, two separate levels of education are distinguished. High-skill migrants include those with at least one year of 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.[4]

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 42.5 million migrants in 1990. About 18.5 million of the 20.9 million high-skill migrants in the OECD countries are concentrated in only six destination countries: the US (10.3 million), Canada (2.7 million), Australia (1.6 million), Israel (1.5 million), the United Kingdom (1.2 million), and Germany (1.2 million).

(b) *Migration data for non-OECD countries*

We further supplement our expanded data collection of our 34 OECD destinations with 66 non-OECD countries in 2000 and 27 countries in 1990, adhering to the same methodological principles and definitions as in DLM. 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 7. In 13 cases, data are obtained directly from the relevant destination countries' national statistical offices. In 23 cases, data are taken from IPUMS-International or the United Nations' Economic Commission for Latin America and the Caribbean (ECLAC) databases, two of the largest archives of publicly available census samples. They are based on samples of at least 5% of the whole population. Data for the six Gulf Cooperation Council (GCC) countries are estimated on the basis of their Labor Force Surveys. The data for the remaining countries were obtained from the OECD DIOC-E database.[5]

When constructing such a large-scale data set, we have to deal with inevitable gaps in the data. Imputation and adjustment are preferred when census data are not sufficiently detailed to identify the gender and education characteristics of stock of migrants in particular bilateral corridors. We therefore distinguish between two types of data sources on the basis of the degree of imputation needed to construct our database.

First, we obtained highly detailed census data from 17 non-OECD countries in 1990 and 60 non-OECD countries in 2000. These countries are not marked with an asterisk in Table 7. They require two minimal adjustments: (i) we use OPSW data to split the available data if there are geographic aggregations for certain regions of origin (e.g., South Asia) or else countries that no longer exist (e.g., ex-Czechoslovakia)[6]; (ii) when the year of census differs from 1990 or 2000 by 4 years or more, which is not atypical since censuses are conducted over decennial cycles. We rescale data using the annual growth rate of the total immigrant stock provided by the United Nations database.[7] Similar adjustments were used in DLM for OECD destinations. We consider these adjustments as minor corrections.

For the remaining 10 non-OECD countries in 1990 and six countries in 2000, we obtained precise data by country of origin, which unfortunately lacked the education and/or gender distribution of immigrants, which we subsequently imputed. These countries are highlighted by an asterisk in Table 7. For four non-OECD countries in 1990 (Bulgaria, Cyprus, Latvia, and Malta), we assume the same gender structure as in 2000 and adjust the education structure proportionately to the changes in educational attainment within the total resident population during 1990–2000. Furthermore, two countries that are now in the OECD, required the same adjustment in 1990 (Estonia and Slovenia). For five Persian Gulf countries (Bahrain, Kuwait, Oman, Qatar, and United Arab Emirates), we apply the gender and skill structure observed in Saudi Arabia. In the case of Saudi Arabia, the education structure was only available for the total immigrant stock and we assumed it is identical across origin countries. For this reason, we also included Saudi Arabia in the group of countries requiring partial imputation.

Adding the 66 non-OECD destination countries increases the overall migrant stock by 35.9 million in 2000, including 5.4 high-skilled migrants and 16.3 million women (see Table 1). The proportion of college graduates among the observed non-OECD countries is 15% and the share of women is 45%, far below the ratios observed in OECD destination countries in both of these dimensions (35% and 51%, respectively). These ratios vary considerably across countries and this heterogeneity is explored in more detail in Section 5. Five of these 66 additional destination countries are home to more than one million foreign-born adults in 2000. These are Cote d'Ivoire (3.9 million), Saudi Arabia (3.1 million), Hong Kong SAR (1.9 million), the United Arab Emirates (1.2 million), and Malaysia (1.0 million).

3. ECONOMETRIC STRATEGY FOR IMPUTING MISSING DATA

Despite the additional data for 66 non-OECD destinations in 2000 and 27 non-OECD destinations in 1990, our bilateral migration matrices remain incomplete. It is important to

emphasize that despite lacking data for a fairly large number of destination countries, the raw data that we have collected nevertheless comprise around 85% of the total in 2000 and two-thirds of the total in 1990. The second major contribution of our paper is in imputing, to the greatest extent possible, those cells for which data are still missing, along both the gender and education dimensions. To this end, we develop a three-step estimation procedure, based upon the most up-to-date theoretical and empirical advances in the literature. We subsequently use the resulting parameter estimates from our empirical exercise to predict the bilateral migrant stocks for cells for which we lack data. While no doubt second-best, given the wider paucity of migration data, we deem our methodology worthwhile in the sense that our results, especially when aggregated over regions, still provide superior estimates of the global winners and losers in the global contest for high-skilled migrants (in a static sense) as when compared to their total absence. [8] A fair comparison to our intuitive approach, which is widely accepted in the literature, are global GDP figures, balance of payments components, and international capital flow estimates, which for many countries are computed in light of the paucity of better quality data.

The data set of OPSW plays a key role in our imputation strategy. While not providing data disaggregated by skill level and additionally comprising migrants of all ages (as opposed to only those aged 25 and above), OPSW embodies significant informational content about the composition of overall bilateral migration stocks globally; arguably far more than any estimation procedure alone could be reasonably expected to capture. These data, which span 1960–2000, prove useful on two counts. First, they provide information on past (pre-1990) migrant stocks, i.e., migrant networks, that we use in our estimations. Second, since the data that we are imputing are a sub-sample of these overall migrant stocks, the OPSW data importantly provide upper-bounds to our estimates that can then be subsequently disaggregated according to this paper's main focus, migrants' level of human capital.

In the next section, we first outline the pseudo-gravity model, a suitable econometric specification for our purposes and further highlight the need for our three-step econometric procedure. We then discuss a number of estimation issues that further need be considered before continuing to an analysis of the accuracy of our results.

(a) *The three step econometric model*

The econometric model that we use to construct our out-of-sample predictions is an extension of the recent developments in the literature. Our theoretical foundation is the random utility model of migration that has been used extensively in the literature (Anderson, 2011; Beine, Docquier, & Ozden, 2011; Grogger & Hanson, 2011; Ortega & Peri, 2012; Beine & Salomone, 2013; Bertoli & Fernandez-Huertas Moraga, 2013). The premise is that individuals with different levels of education are assumed to choose between staying at home or else moving to alternate destinations; with their decisions based upon the utility or income they are expected to receive in competing destinations as when compared to remaining sedentary. Their utility is the sum of a deterministic component (capturing dyadic migration costs as well as origin and destination-specific push and pull factors) and a random term (capturing individual heterogeneity in migration tastes). We assume that the random term is iid and extreme-value distributed, which implies that the ratio of bilateral migrants to stayers only depends upon destination and origin country characteristics. [9]

Our model combines the approaches of Beine *et al.* (2011) and Grogger and Hanson (2011). Following Grogger and Hanson (2011), our dependent variable is the gender and education-specific bilateral stock of migrants and similarly to Beine *et al.* (2011) we allow these stocks to depend upon the lagged level of the total diaspora. This dynamic specification allows us to account for any inertia in the evolution of migration stocks and for the attractive power of existing diasporas. We use the following specification:

$$M^{jk}_{g,s,t} = \exp\left(\alpha_{0,g,s,t} + \alpha_{1,g,s,t} d^{jk}_{g,t} + \alpha_{2,g,s,t} b^{jk} + \gamma^{j}_{g,s,t} + \lambda^{k}_{g,s,t}\right) + \epsilon^{jk}_{1,g,s,t}, \quad (1)$$

The variables are defined as follows. [10] The dependent variable $M^{jk}_{g,s,t}$ is the bilateral stock of migrants from country j in country k in year t (either 1990 or 2000), of gender g and skill (education) s. The explanatory variables comprise an historical bilateral time-varying diaspora variable, $d^{jk}_{g,t}$, which is a key determinant of future migration levels Beine *et al.* (2011). In addition, we have various time-invariant bilateral variables, denoted by b^{jk}, such as geographic distance, common language, contiguous borders, and shared colonial heritage that account for cultural, political, and geographic linkages. In this ideal set-up, origin country characteristics (such as economic, political and social push factors) are captured by a set of origin fixed effects introduced through the vector $\gamma^{j}_{g,s,t}$. Similarly comparable (pull) factors at destination would be accounted for by the inclusion of vector $\lambda^{k}_{g,s,t}$ of destination fixed effects. Since our goal is to impute migration data for those cells for which destination data are missing, it is not possible to include vector $\lambda^{k}_{g,s,t}$ in our regressions however. This gives rise to our *prediction problem*, which our three-step estimation seeks to address. [11]

Our preliminary step is to run a first-stage gravity regression with education-aggregated migration data disaggregated by gender (but not skill level), obtained from OPSW on the left hand side, to recover estimates of the gender-specific pull variable $\lambda^{k}_{g,t}$ (i.e., destination fixed effects aggregated over the education dimension) for all countries in our sample. This takes the form:

$$M^{jk}_{g,t} = \exp\left(\alpha_{5,g,t} + \alpha_{6,g,t} d^{jk}_{g,t} + \alpha_{7,g,t} b^{jk} + \lambda^{k}_{g,t} + \theta^{j}_{g,t}\right) + \epsilon^{jk}_{2,g,t}. \quad (2)$$

In this equation, $M^{jk}_{g,t}$ is observed for all origins and destinations in OPSW and it is equal to the sum of all education-specific bilateral migrant stocks for a given gender and time period. In line with Eqn. (1) we include a comprehensive set of dyadic variables to capture migration costs as well as origin and destination fixed effects in each regression. The estimation of Eqn. (2) thus furnishes us with estimates of the time and gender-specific destination pull variables, $\hat{\lambda}^{k}_{g,t}$.

Since the goal of our initial econometric procedure is to compute estimates of gender and education-specific destination fixed effects, we further parameterize the gender–education destination pull factor as:

$$\lambda^{k}_{g,s,t} = \alpha_{3,g,s} A^{k}_{t} + \alpha_{4,g,s} \lambda^{k}_{g,t} \quad (3)$$

In this expression, $\lambda^{k}_{g,t}$ is the gender-specific pull variable and A^{k}_{t} is a vector of destination-specific parameters, included in estimation to further distinguish across skill groups. These include (for the destination country) whether people speak English, the size of the total labor force (in logs), GDP per capita (in logs), the total fertility rate (in logs), the ratio of the number of highly skilled to the total labor force, and the

labor force participation rates of both men and women. A number of dummy variables are also included that capture whether a destination country belongs to the GCC, whether military service is compulsory and whether polygamy is legally practiced. [12]

Putting everything together, we substitute Eqn. (3) into Eqn. (1), which yields our second-stage regression:

$$M^{jk}_{g,s,t} = \exp\left(\alpha_{0,g,s,t} + \alpha_{1,g,s,t} d^{jk}_{g,t} + \alpha_{2,g,s,t} b^{jk} + \alpha_{3,g,s} A^k_t + \alpha_{4,g,s} \hat{\lambda}^k_{g,t} + \gamma^j_{g,s,t}\right) + \epsilon^{jk}_{1,g,s,t}. \quad (4)$$

When compared with Eqn. (1), A^k_t is the vector of destination-specific parameters from Eqn. (3). The gender-specific destination pull variables, $\hat{\lambda}^k_{g,t}$ are those estimates obtained from our first-stage regression, Eqn. (2). Our two-step procedure up until this point, which aims to maximize the accuracy of our predictions, comes at a cost; our gender-specific destination pull variables, $\hat{\lambda}^k_{g,t}$, will be correlated with the vector A^k_t of destination-specific variables. In turn, this means that it is impossible to meaningfully interpret A^k_t and as such these results are not reported. Given that more data exist in 2000, when compared to 1990 and again, so as to maximize the accuracy of our predictions, for each gender–education pair we run regressions for both 1990 and 2000 simultaneously, allowing our explanatory variables to vary over time, but additionally forcing the coefficients on these variables to be constant in both years.

Until this point, we have four separate gender–education pairs: high-skilled men, high-skilled women, low-skilled men, and low-skilled women. In our final step, we take our predictions for the missing cells from Eqn. (4) and apply the resulting shares to the totals detailed in OPSW. Since these data provide the most accurate estimates of global migrant stocks, we are, in effect, splitting the OPSW data into the relevant education–gender bilateral migrant stocks. Herein lies the final econometric issue that needs addressing, namely that OPSW refers to migrants of all ages, while DLM instead only refers to those in the labor force. The difference are those aged 24 and below, which for the sake of simplicity we term *youth*. To surmount this final issue, we run estimate Eqn. (4) twice more, for male youth and female youth. [13] Finally, putting together all of our estimates, for those destination countries for which raw migration data are unavailable we lastly use the following restriction to split the total gender-specific migrant stock provided by OPSW into the gender–skill-specific migrant stock:

$$\widetilde{M}^{jk}_{g,s,t} = \frac{\widehat{M}^{jk}_{g,s,t}}{\sum_s \widehat{M}^{jk}_{g,s,t}} M^{jk}_{g,t}. \quad (5)$$

Our final matrices are comprised of the original gender–education-specific data for those destination countries for which we have original raw data in together with our predicted migrants stocks, $\bar{M}^{jk}_{g,s,t}$.

(b) *Econometric issues*

The presence of a large number of zero or undefined observations in the dependent variable (gender- and education-specific bilateral migrant stocks) gives rise to econometric concerns that would yield inconsistent OLS estimates. Zero observations appear in large numbers in many bilateral contexts such as international trade, official aid, military conflict, and political alliances. This phenomenon is especially prevalent in migration data sets, since there is no observed or recorded migration between many country pairs, for example, between Rwanda and Mongolia, due to high geographic, cultural, and economic barriers. Furthermore, censuses or alternative surveying instruments are unlikely to capture small migration corridors should any sampling strategy be followed. As a result, we observe zero values for about 48.5% of the 18,900 observations ($(190 - 1) = 189$ destination $\times$ 100 origin countries) in the aggregate migration matrix from OPSW for 2000. The ratio of zero observations is 52.6% for low-skilled males, 52.9% for high-skilled males, 52.8% for low-skilled females, and 54.0% for high-skilled females. [14]

Two main reasons explain why a high proportion of zero observations in the dependent variable typically results in inconsistent parameter 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 with the occurrence of non-random zero observed flows. In a double log regression model, the norm therefore is to add "one" to the value of the dependent variable and then take the log. 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 heteroscedasticity, 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 Likelihood (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. Actually, our discrete choice specification is a special case of Artuc, Artuc, Chaudhuri, and McLaren (2010) self selection model. Artuc (2013) provides a detailed comparison of PPML and its alternatives for estimating relevant discrete choice models. In the PPML regression robust standard errors are implemented.

(c) *Estimation results*

The results from our estimation are presented in Tables 2a and 2b. The first table is for the first stage, Eqn. (2), which generates the estimates of the time and gender-specific destination pull variables, $\hat{\lambda}^k_{g,t}$. In addition, the estimation generates coefficients for the bilateral variables that are gender specific but are not skill specific. These coefficients all have the expected signs and are consistent with the results from the literature.

The more interesting results are those of the determinants of our four gender–education-specific migration pairs estimated simultaneously for 1990 and 2000, which are presented in Table 2b. As noted in the previous section, we only report our estimates of the bilateral variables since the parametrization of our gender-specific destination (pull) fixed effect $\lambda^k_{g,t}$ means that the coefficients on our destination-specific variables, A^k_t are not interpretable. All the bilateral variables that capture various aspects of migration costs are highly significant at the 1% level, with sensible orders of magnitude relative to previous findings in the literature, which moreover vary sensibly over both skill groups. Geographical distance deters international migration and has a stronger impact upon the low-skilled, since typically high-skilled migrants are better able

Table 2a. *First stage regression results*

	Female		Male	
	Year 1990	Year 2000	Year 1990	Year 2000
Language	0.483[a]	0.407[a]	0.373[a]	0.279[a]
	(0.042)	*(0.035)*	*(0.038)*	*(0.033)*
Border	0.401[a]	0.393[a]	0.407[a]	0.425[a]
	(0.045)	*(0.037)*	*(0.043)*	*(0.036)*
Distance	−0.488[a]	−0.334[a]	−0.492[a]	−0.344[a]
	(0.021)	*(0.018)*	*(0.021)*	*(0.018)*
Colonial link	0.660[a]	0.029	0.627[a]	0.058
	(0.045)	*(0.041)*	*(0.045)*	*(0.041)*
Diaspora	0.457[a]	0.550[a]	0.467[a]	0.553[a]
	(0.007)	*(0.007)*	*(0.006)*	*(0.006)*
OBS	35,910	35,910	35,910	35,910
RSQR	0.908	0.915	0.847	0.885

[a] Significant at 1% level.

Table 2b. *Second stage regression results*

	Female				Male			
	High-skill		Low-skill		High-skill		Low-skill	
	Year 2000	Year 1990	Year 2000	Year 1990	Year 2000	Year 1990	Year 2000	Year 1990
Language	0.663	0.553	0.484	0.454	0.577	0.518	0.153	0.258
	(0.030)	*(0.045)*	*(0.040)*	*(0.049)*	*(0.030)*	*(0.042)*	*(0.042)*	*(0.051)*
Border	0.281	0.223	0.808	0.606	0.515	0.222	0.844	0.488
	(0.036)	*(0.056)*	*(0.043)*	*(0.057)*	*(0.040)*	*(0.059)*	*(0.047)*	*(0.062)*
Distance	−0.247	−0.348	−0.381	−0.524	−0.163	−0.280	−0.395	−0.585
	(0.013)	*(0.017)*	*(0.019)*	*(0.021)*	*(0.014)*	*(0.017)*	*(0.021)*	*(0.023)*
Colonial link	0.496	0.708	0.211	0.644	0.498	0.588	0.349	0.653
	(0.030)	*(0.044)*	*(0.041)*	*(0.051)*	*(0.034)*	*(0.049)*	*(0.047)*	*(0.058)*
Diaspora	0.394	0.342	0.537	0.463	0.419	0.387	0.542	0.495
	(0.005)	*(0.006)*	*(0.007)*	*(0.008)*	*(0.005)*	*(0.007)*	*(0.007)*	*(0.008)*
Destination fixed effect	0.911	0.911	1.084	1.084	0.989	0.989	1.148	1.148
	(0.016)	*(0.016)*	*(0.020)*	*(0.020)*	*(0.018)*	*(0.018)*	*(0.022)*	*(0.022)*
OBS	30,419		30,419		30,419		30,419	
RSQR	0.893		0.898		0.871		0.871	

Notes. All estimates are significant at 1% level.

to overcome higher international migration costs. Similarly, while migrants from both skill groups migrate more on average to bordering countries, this effect is much stronger for the low-skilled for similar reasons. Furthermore, both of these effects are similar in magnitude in both 1990 and 2000. Sharing a common language is more important for the high-skilled since language requirements are generally more relevant for this group. Diaspora networks conversely encourage the migration of low-skilled migrants more than their high-skilled counterparts, since this latter grouping typically has additional resources in order to migrate internationally. An alternative explanation is that networks will likely play a more important role in South-South migration, which comprises higher volumes of low-skilled migrants. Similarly, colonial links, which may be considered to also comprise an historical diaspora component are more relevant for the low-skilled, but across both skill groups the influence of colonial links declines between 2000 and 1990, a result in line with Head, Mayer, and Ries (2010). All of these results on the role of the bilateral variables are consistent with other papers in the literature, such as Beine *et al.* (2011) and Grogger and Hanson (2011). The differences in the magnitudes of the parameter estimates are likely driven by the fact that previous papers used data from OECD destinations whereas we are able to include many non-OECD destinations.

Once the estimation is complete, we use the estimated parameter values from the equations above to predict the missing values in the bilateral migration matrices for both years, genders, and education levels.

One critical issue is that of zero observations where two basic options are available. The first would be to use PPML estimates. It is worth noting, however, that our predicted numbers are almost never integers and so our estimates should be interpreted as the expected numbers of migrants. Although PPML predictions cannot produce true zeroes, many of our estimates are extremely close to zero, such that we could postulate having predicted a zero, e.g., if the expected number of migrants generated by the PPML model is smaller than one (or indeed below any other arbitrarily small threshold). The second option which we implement in this paper, instead involves splitting the aggregate figures provided by OPSW by using our PPML regressions to estimate the shares of males and females and the skilled and unskilled. Since we have access to the totals, predicting these shares is no doubt more accurate than attempting to impute unrestricted totals. In following this methodology, we necessarily take all zero observations from OPSW as opposed to imputing them with our PPML estimation. Indeed, there are a very few cases where OPSW is greater than zero but where a PPML estimate would be close is zero, but these observations constitute are less than 1% of the total.

Before analyzing the predictions and their implications for the global migration patterns by skill level, we evaluate the out-of-sample performance of our predictions. To do so, we first randomly drop low-skilled and high-skilled migrant stock data for five countries from the sample for which we have actual data. We then estimate the gravity model as explained in the previous section, as if these five destination countries also had missing data. We impute the missing low-skilled and high-skilled migrant stocks for these five countries and calculate the log-ratios of these imputed data to the actual migrant data. We repeat these four steps 100 times and subsequently evaluate their performance.

Figure 2 presents the density functions of these log-ratios for female migrants for 2000. The left-hand side graphs refer to low-skilled migrants, while the right-hand side figures refer to high-skilled migrants. We use different cutoffs to assess the predictions of corridors of different sizes. More specifically, the top graphs capture all corridors, the middle set refers to corridors comprising greater than 1,000 migrants, while the bottom graphs refer to corridors of greater than 10,000 migrants. In each case, the densities are bell-shaped, and the median is around zero. [15] Inherent idiosyncratic factors exist in the estimation of small migration corridors. As clearly seen in the comparison of Figure 2a–c or d–f, the distribution of smaller corridors (a and d) span a wider range, indicating a larger standard deviation. As corridors get larger however, the distribution of the log ratio becomes increasingly centered around zero, indicating higher precision. Since these larger corridors comprise the vast majority of the migrant stocks, their higher precision increases our confidence in overall interpretation of our results.

4. A REFINEMENT OF GLOBAL BRAIN DRAIN INDICATORS

Our raw data in combination with our imputed data allow us to improve upon existing high-skilled migration indicators (e.g., Docquier & Marfouk, 2006; Docquier *et al.*, 2009; Dumont & Lemaître, 2004; Dumont *et al.*, 2007).

Previous studies, focusing upon the subset of OECD destinations, 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. [16] 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 countries.

First we define the following key variables:

$M^{jk}_{g,s,t}$: the stock of bilateral migrants from country j to country k of gender g and skill s at time t

$I^{i}_{g,s,t}$: the stock of total immigrants of type (g,s) to country i in year t

$E^{i}_{g,s,t}$: the stock of total emigrants of type (g,s) from country i in year t

$L^{i}_{g,s,t}$: the (observed) resident labor force of type (g,s) in country i in year t

$N^{i}_{g,s,t}$: the natural labor force of type (g,s) in country i in year t. This is the number of workers from a given country i regardless of their current location.

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

$$I^{i}_{g,s,t} \equiv \sum_{j} M^{ji}_{g,s,t} \tag{6a}$$

$$E^{i}_{g,s,t} \equiv \sum_{k} M^{ik}_{g,s,t}. \tag{6b}$$

We then use data on both 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 (natives residing in their country of birth) plus immigrants. Similarly, the natural labor force of type (g,s) in country i, $N^{i}_{g,s,t}$, is equal to the *non-migrant* labor force plus emigrants. We can therefore state that the non-migrant labor force can be expressed as either of the following expressions (residents minus immigrants or naturals minus emigrants):

$$L^{i}_{g,s,t} - I^{i}_{g,s,t} = N^{i}_{g,s,t} - E^{i}_{g,s,t} \tag{7}$$

The ability to recover our measure of the natural labor force $N^{i}_{g,s,t}$, a prerequisite for which is to have measures of immigrant/emigrant stocks for all countries 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 construct $N^{i}_{g,s,t}$ using Eqn. (7), we first need to construct a consistent measure of $L^{i}_{g,s,t}$, i.e., the resident labor force.

We begin with a measure of the total working-age population (i.e., aged 25 and over) by gender as provided by the United Nations. [17] Data are missing for a few countries and these are instead obtained from the CIA World Factbook. [18] 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 *et al.* (2009) and combine different data sets documenting the proportion of post-secondary educated workers in the population aged 25 and over (i.e., Barro & Lee, 2001; Doménech & de la Fuente, 2006; Cohen & 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 sometimes less than 1%. Given the construction of $L^{i}_{g,s,t}$, Eqn. (7) 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.

With all the constituent components in hand we now define 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:

$$e^{i}_{g,s,t} \equiv \frac{E^{i}_{g,s,t}}{N^{i}_{g,s,t}}, \quad b^{i}_{g,s,t} \equiv \frac{E^{i}_{g,s,t} - I^{i}_{g,s,t}}{N^{i}_{g,s,t}} \tag{8}$$

so that (7) 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 quantify 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 111.9 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

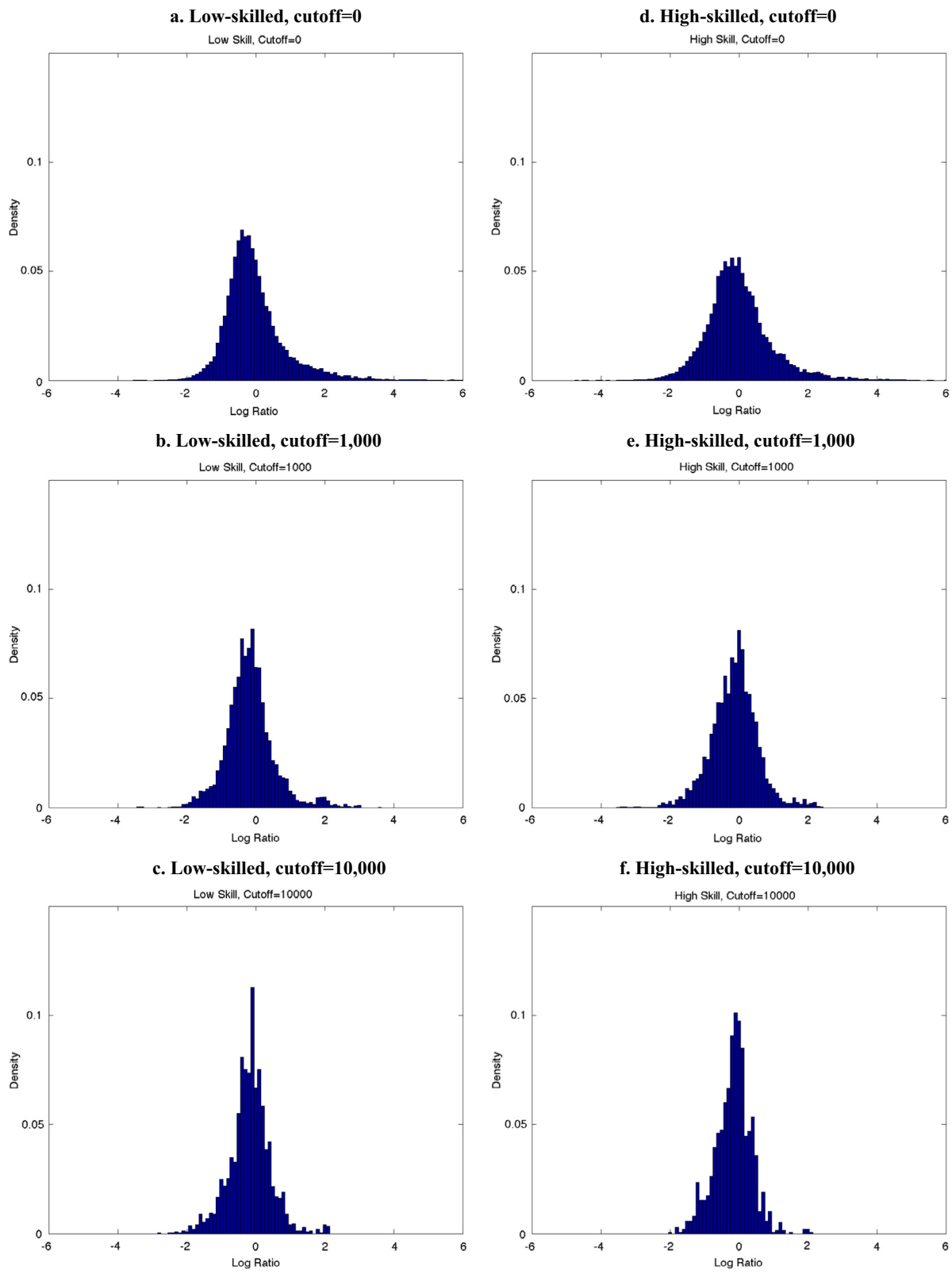


Figure 2. *Histograms of out-of-sample log ratios, Female migrants, 2000.*

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 at the higher skill level.

Net vs Gross—We are able to identify the size and skill structure of adult immigration and emigration stocks from all countries, including for those in the developing world. In

previous studies, immigration data were only available for OECD member states and selected non-OECD countries. Complete emigration data were simply unavailable. Furthermore, data only referred to gross immigration. Since we have both immigration and emigration numbers for all countries, we can compare the entries and exits of workers and compute comparable net migration balances for college graduates and less educated workers for all countries globally.

Given the breadth of our data set, we are also able to characterize the skill levels of the natural population. Given (7) and (8), average skill levels of naturals and residents are linked through the following equation, where the subscripts h and l stand for the high and low skilled, respectively:

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

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 of positive selection in international migration ($e^i_{h,s,t} > e^i_{l,s,t}$). However, what matters are the net emigration rates of high-skilled and low-skilled workers. International migration reinforces human capital inequalities across countries if $b^i_{k,h,t} > b^i_{k,l,t}$. We 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.

Table 3. *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 (million)	College (%)	Women (%)
Year 2000									
WORLD	111.9	25.7	48.7	59.3	35.3	51.0	52.6	15.0	46.1
OECD	32.2	30.4	50.4	29.1	31.0	50.8	3.1	24.4	46.6
HIGH	26.3	36.0	52.0	22.3	38.4	53.0	4.0	22.6	46.6
DEV	85.6	22.6	47.7	37.0	33.4	49.8	48.6	14.3	46.1
LOW	15.5	9.6	45.0	2.5	38.0	48.5	13.0	4.1	44.3
LDC	15.1	8.5	43.6	2.4	34.6	47.7	12.7	3.5	42.8
SIDS	4.3	35.6	54.6	4.0	37.0	54.9	0.3	17.7	51.7
USA	0.9	58.7	50.0	0.7	62.9	52.6	0.2	45.0	41.4
CANZ	1.5	57.1	54.0	1.4	57.6	54.3	0.1	46.9	47.9
EU27	20.0	32.3	52.0	17.7	33.1	52.4	2.3	25.6	48.8
GCC	0.6	20.3	37.3	0.0	65.2	39.7	0.6	16.5	37.0
LAC	15.6	25.1	50.2	14.0	26.4	50.1	1.6	13.9	51.6
SSA	10.5	11.5	45.2	2.2	43.1	47.5	8.3	3.1	44.6
CIS	19.2	26.3	54.6	2.4	42.1	58.2	16.8	24.0	54.1
INDIA	6.1	23.7	36.9	1.7	60.5	47.2	4.4	9.4	32.9
CHINA	3.9	27.9	51.8	1.7	46.7	53.0	2.3	14.0	51.0
MENA	9.1	21.8	37.9	4.2	29.9	43.0	4.9	14.7	33.5
Year 1990									
WORLD	85.3	19.1	47.9	42.5	29.5	50.7	42.7	8.7	45.2
OECD	25.7	26.2	51.1	23.3	26.9	51.6	2.4	19.4	46.3
HIGH	23.4	29.1	52.0	20.5	30.5	52.8	2.9	18.9	46.4
DEV	61.9	15.3	46.4	22.1	28.6	48.7	39.8	7.9	45.1
LOW	13.1	7.2	42.9	1.4	33.7	45.6	11.7	4.0	42.6
LDC	13.0	6.5	41.4	1.4	30.2	45.1	11.6	3.7	40.9
SIDS	3.0	31.0	52.6	2.6	34.6	53.6	0.4	7.8	46.5
USA	0.8	51.4	50.4	0.6	53.8	53.0	0.2	43.3	41.3
CANZ	1.3	46.0	56.1	1.2	46.4	56.3	0.1	38.2	52.6
EU27	18.8	25.3	51.9	16.9	26.0	52.2	1.9	19.4	49.2
GCC	0.4	17.7	34.2	0.0	64.8	35.6	0.4	14.3	34.1
LAC	8.2	24.7	50.4	7.0	27.4	50.7	1.2	9.3	48.2
SSA	8.5	7.5	44.1	1.2	39.6	44.3	7.3	2.1	44.1
CIS	14.1	12.7	57.4	1.8	20.8	56.3	12.2	11.5	57.6
INDIA	5.5	12.4	35.1	1.0	45.5	47.0	4.5	5.2	32.5
CHINA	3.3	16.5	50.7	0.9	40.0	50.2	2.4	7.8	50.9
MENA	6.8	17.3	36.1	3.2	23.8	41.5	3.6	11.6	31.3

Notes. Column "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.

5. A GLOBAL ASSESSMENT OF HUMAN CAPITAL MOBILITY

In this section, we first provide some overarching descriptive statistics in order to highlight the global patterns of international migration in 1990 and 2000. We then study country-specific characteristics and identify the main source countries, focusing upon college-graduate migrants (Section (a)) and high-skilled female migrants (Section (b)). Finally, we will compare the concepts of human capital per resident and per natural (Section (c)).

Table 3 details total emigrant stocks and their education/gender composition in 1990 and 2000 for key regions or income categories of the world. The top portion of Table 3 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 the following geographical regions: (1) the United States, (2) Canada, Australia, and New Zealand as a single entity, which is referred to as CANZ, (3) the 27 countries 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 numbers reveal that as income levels increase so do the percentages of high-skilled emigrants and female emigrants abroad. Comparing emigrations from these regional groupings to OECD and non-OECD

Table 4. *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	8.1	5.9	2.2	0.0	0.0	0.0
OECD	4.8	4.4	0.4	−5.5	−5.4	−5.7
HIGH	4.8	4.4	0.5	−6.6	−6.8	−6.4
DEV	12.0	7.7	4.3	8.0	7.0	9.6
LOW	20.3	13.0	7.3	16.0	13.7	21.2
LDC	19.9	12.9	7.0	16.6	14.9	21.0
SIDS	40.9	39.3	1.6	34.8	29.3	40.8
USA	0.6	0.5	0.1	−11.6	−12.0	−11.2
CANZ	7.2	6.9	0.2	−30.9	−32.4	−29.5
EU27	9.7	8.8	0.9	2.4	2.4	2.3
GCC	14.3	3.6	10.7	−104.9	−230.7	−32.2
LAC	12.2	11.5	0.7	10.8	9.8	11.9
SSA	15.6	12.3	3.3	10.2	8.6	13.7
CIS	16.1	3.2	12.9	2.8	2.2	3.6
INDIA	6.0	4.3	1.7	5.4	4.7	7.0
CHINA	5.2	3.7	1.5	5.2	3.7	9.2
MENA	17.5	11.2	6.4	9.0	10.3	6.6
Year 1990						
WORLD	6.8	5.2	1.5	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	−4.7	−4.7	−4.7
DEV	9.6	6.4	3.2	6.8	6.0	8.1
LOW	23.2	11.8	11.4	18.9	17.7	21.8
LDC	23.2	11.7	11.6	20.4	19.2	24.1
SIDS	42.2	40.7	1.4	38.0	33.0	43.8
USA	0.7	0.5	0.1	−10.0	−9.4	−10.8
CANZ	6.6	6.3	0.3	−28.8	−30.5	−27.0
EU27	9.2	8.5	0.7	3.9	3.8	4.1
GCC	12.9	3.1	9.8	−104.1	−188.2	−36.4
LAC	10.7	10.1	0.6	9.3	8.3	10.5
SSA	17.2	13.1	4.1	10.6	9.0	15.2
CIS	7.6	1.6	6.0	1.8	1.6	2.1
INDIA	4.3	2.8	1.5	2.6	2.1	3.9
CHINA	4.5	3.0	1.6	4.5	3.2	10.5
MENA	19.9	12.9	7.0	13.7	13.9	13.2

Notes. Column "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.

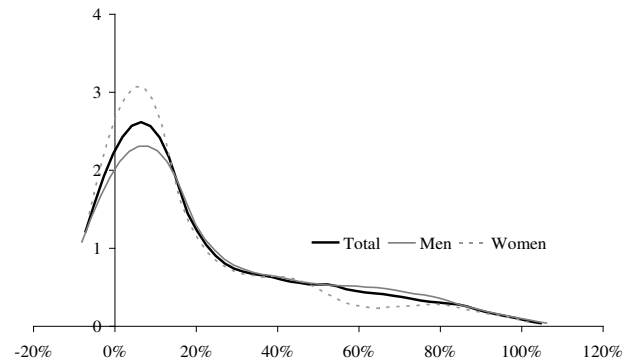
destinations further reveals the strong selection inherent in world migration patterns. Across all regional groups, a far higher proportion of both college educated and women emigrate to OECD destinations. This selection on skills is most pronounced in the cases of low income and least developed countries from which only 4.1% and 3.5% of emigrants to non-OECD countries have college education as opposed to 38.0% and 34.6% in OECD countries respectively. These patterns are also reflected strongly in the data for 1990.

The second section of the top and bottom panels of Table 3, again reveal strong patterns of selection. The proportions
of both the high-skilled and women emigrants are far larger in OECD destinations when compared to non-OECD destinations in 2000; 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 countries, in terms of high-skill composition increased during 1990–2000, with the exception of those from the GCC and the Commonwealth of Independent States, which over time both sent more highly skilled migrants to other non-OECD destinations. Similarly, although many regions send larger numbers of female migrants abroad in both 1990 and 2000 e.g., the GCC, Latin America, and the Caribbean, Sub-Saharan Africa, the Commonwealth of Independent States, India, and China, the selection on females increasingly favored the OECD from all these regions with the exception of the GCC and the Commonwealth of Independent States.

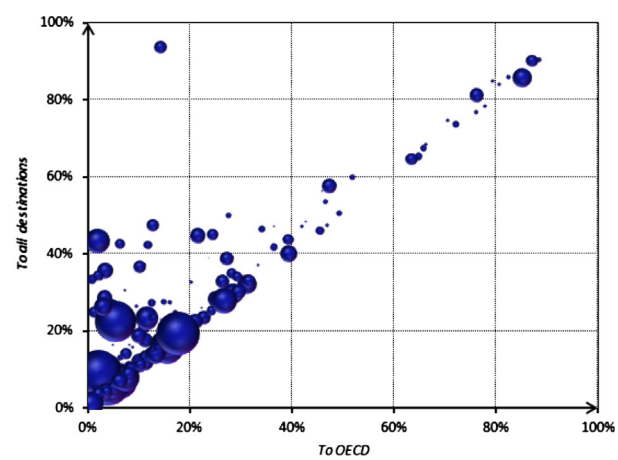
Columns 1 and 4 in Table 4, provide gross and net emigration rates, calculated according to Eqn. (4). For gross rates, we further distinguish between emigration to OECD and non-OECD countries (columns 2 and 3). Net rates are provided for men and women with college education (columns 5 and 6). Globally, 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 19.9%, respectively. The most affected geographic regions are the MENA (17.5%), CIS (16.1%), Sub-Saharan Africa (15.6%), and the GCC (14.3%). 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 the ex-Soviet block, the GCC, and MENA regions.

A comparison of gross and net emigration rates proves highly instructive. High-income and OECD countries exhibit negative net high-skilled migration rates i.e., the incoming pool of educated people to those regions more than compensates for any human capital loss suffered as a consequence of their high-skilled nationals emigrating abroad. Consequently, international high-skilled mobility increases the number of college graduate workers in the labor force by over 10% in the United States, around 30% in other settlement countries (Canada, Australia, and New Zealand) and remarkably more than doubles this proportion in the oil-producing countries of the GCC. As 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

a. Density of "non-OECD to total" ratio of emigration rates



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



c. Net versus gross emigration rates in 2000

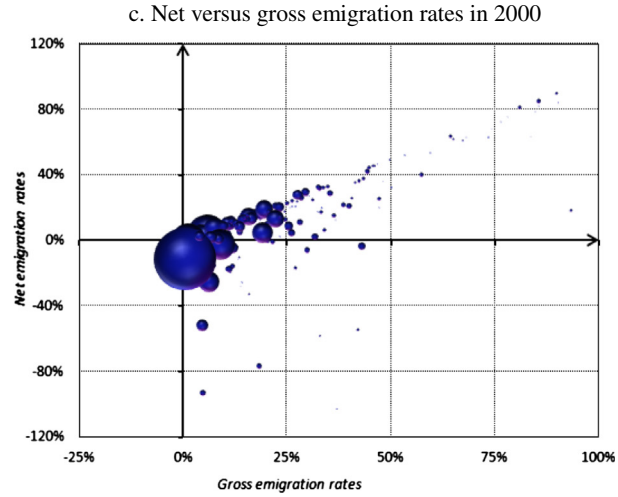


Figure 3. *Distribution of high-skilled emigration rates. Notes. On Figures b and c. each country is a represented by a bubble, the size of which is proportional to the high-skilled emigration stock.*

Table 5. *Brain drain: most and least affected countries*

Largest high-skilled net emigration rates							Lowest high-skilled net emigration rates						
Country	2000			1990			Country	2000			1990		
	Net (%)	Gross (%)	Non-OECD (%)	Net (%)	Gross (%)	Non-OECD (%)		Net (%)	Gross (%)	Non-OECD (%)	Net (%)	Gross (%)	Non-OECD (%)
Jamaica	84.6	85.6	0.6	85.5	86.5	0.9	United Arab Emirates	−309.4	9.0	70.6	−103.5	3.0	77.1
Haiti	80.9	81.0	5.9	73.5	73.7	7.9	Saudi Arabia	−93.3	5.0	63.5	−132.4	5.6	61.8
Liberia	53.3	59.8	13.0	55.0	62.9	13.7	Israel	−77.1	18.4	18.2	−19.4	13.2	14.6
Sierra Leone	51.9	53.4	12.7	46.7	48.5	14.0	Oman	−58.7	33.1	98.2	−55.0	29.2	98.4
Eritrea	49.2	49.9	44.6	48.1	49.7	48.6	Kuwait	−54.6	42.2	71.8	−39.4	37.9	75.6
Laos	45.5	45.8	0.5	40.3	42.0	1.8	Australia	−51.9	4.7	8.9	−47.9	3.4	8.0
Somalia	45.1	46.2	26.0	32.2	34.0	30.8	Canada	−25.8	6.4	2.8	−22.9	6.5	3.5
Afghanistan	44.1	45.0	45.5	25.8	26.9	59.4	Switzerland	−17.7	11.2	8.0	−12.4	8.2	11.0
Bosnia and Herzegovina	41.7	44.6	51.5	29.7	33.8	29.0	Singapore	−16.0	12.0	19.6	1.0	11.1	13.4
Lebanon	39.8	57.5	17.6	52.3	65.4	16.9	Libya	−15.8	7.7	19.7	−22.0	9.3	14.7
Kenya	37.2	43.5	9.5	48.7	50.5	11.4	United States	−11.6	0.6	18.1	−10.0	0.7	18.9
Yemen	36.0	42.5	84.8	55.4	59.8	73.0	Sweden	−6.8	5.2	2.7	−3.9	4.3	3.9
Uganda	34.9	41.6	12.0	42.5	43.9	19.7	New Zealand	−6.4	30.0	1.3	−19.9	25.1	2.9
Macedonia	32.6	34.9	18.7	26.1	30.3	6.5	Netherlands	−4.9	12.0	6.7	−2.4	12.1	6.4
Sri Lanka	32.4	32.8	19.2	34.0	37.0	33.9	Paraguay	−4.2	6.2	33.8	−4.7	4.6	8.1
Congo, Rep. of the	32.3	33.6	21.0	14.2	22.0	25.6	Cote d'Ivoire	−4.1	12.9	49.8	−16.5	7.5	25.1
El Salvador	31.9	32.9	3.7	32.9	33.7	5.1	Kazakhstan	−3.8	43.2	95.1	−8.8	9.6	90.1
Nicaragua	31.8	33.9	13.1	28.6	30.0	8.1	Russia	−2.7	9.0	75.3	1.0	4.7	75.9
Cuba	29.4	29.5	3.4	31.6	31.8	2.9	Belgium	−2.6	6.7	9.2	−0.8	5.3	6.2
Azerbaijan	28.5	35.5	90.1	−4.0	5.3	75.6	France	−2.5	4.0	13.8	−1.1	3.1	12.1
Vietnam	27.4	27.7	3.1	23.9	24.0	1.1	Spain	−2.4	4.3	15.9	−0.8	3.4	11.6
Chad	26.8	30.4	75.8	25.3	27.3	77.6	Nepal	−1.7	5.9	27.5	21.1	23.6	79.3
Georgia	25.6	28.5	88.1	11.3	16.1	90.6	Latvia	−0.8	20.5	45.5	−34.8	27.1	63.9
Armenia	24.9	47.2	72.7	24.5	29.4	66.2	Germany	0.2	6.8	9.2	2.4	6.8	4.6
Cambodia	24.6	31.4	0.6	26.9	27.8	1.9	Costa Rica	0.2	8.4	8.8	−6.0	10.5	8.5

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.

migration flows. Turning finally to gender differences, the final columns of Table 4, demonstrate that in all regions, net emigration rates are lower for males than for females, with the exception of the EU27 and MENA.

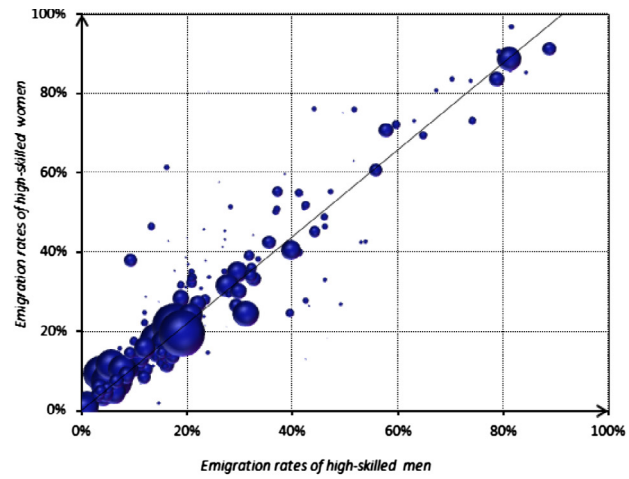
(a) *Country-specific results*

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% (7.9 million over 28.8 in 2000 and 3.7 over 16.3 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.[19] The peak of this kernel density plot corresponds to a ratio of just 0.065 and in the majority of cases (123 out of 190), 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 not less than a sixth of the sample. The individual countries that comprise the thick right-hand tail of the distribution include countries of the Middle-East (that predominantly send emigrants to oil-producing countries), Southern African countries (that principally send migrants to South Africa) and ex-Soviet-block members, which are characterized by significant and voluminous migrations between one another.

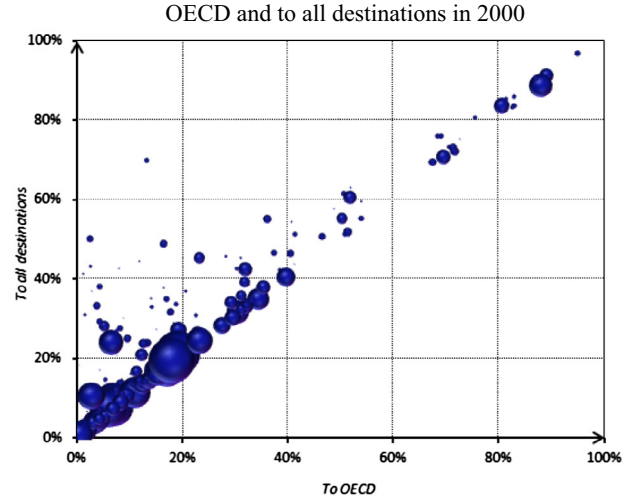
Unsurprisingly, for many countries, a significant disparity exists when comparing high-skilled emigration rates to all destinations as when compared to the OECD alone, which until now has been the focal group in the literature. These marked differences are illustrated in Figure 3b, which plots, for each country the gross emigration rates of college graduates to OECD destinations on the *x*-axis, against those to all destinations on the *y*-axis. In doing so, the figure highlights the importance of our comprehensive global approach. Each bubble in Figure 3b 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 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 points difference for the West Bank and Gaza, 37 percentage points for Yemen, 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 scatter plot 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 countries of the GCC.

a. Emigration rates of high-skilled women and men



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



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

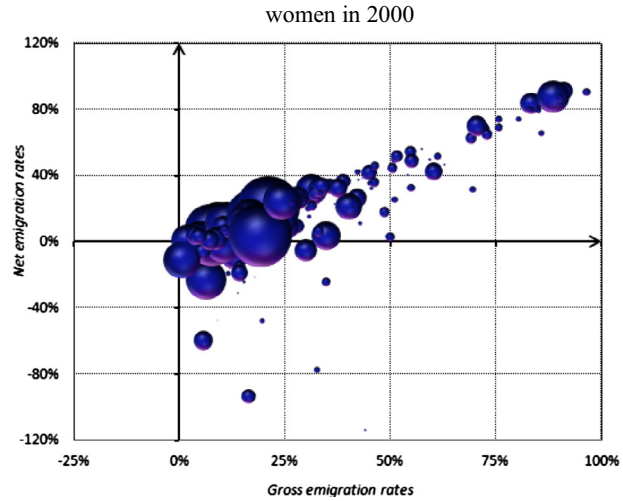


Figure 4. *Distribution of emigration rates of high-skilled women. Notes. On Figures 3b and c. each country is a represented by a bubble. the size of which is proportional to the emigration stock of high-skilled women in 2000.*

Table 6. *Women's brain drain: most and least affected countries*

Highest net high-skilled emigration rates							Lowest net high-skilled emigration rates						
Country	2000			1990			Country	2000			1990		
	Net (%)	Gross (%)	Non-OECD (%)	Net (%)	Gross (%)	Non-OECD (%)		Net (%)	Gross (%)	Non-OECD (%)	Net (%)	Gross (%)	Non-OECD (%)
Jamaica	87.7	88.6	0.7	87.5	88.5	1.0	United Arab Emirates	−202.9	6.6	69.4	−123.4	2.4	69.1
Haiti	83.3	83.5	3.4	78.8	79.0	5.8	Israel	−93.5	16.5	14.2	−20.5	11.2	12.2
Sierra Leone	73.7	75.9	9.6	72.1	74.4	13.5	Kuwait	−77.7	32.9	56.6	−73.3	26.8	64.0
Liberia	68.7	75.8	8.8	64.3	72.5	11.5	Australia	−59.9	5.8	7.6	−66.7	5.2	6.7
Afghanistan	54.3	54.8	34.0	45.8	47.1	51.8	Canada	−23.0	6.4	2.3	−18.8	6.8	3.0
Laos	51.4	51.7	0.5	42.9	45.4	2.1	Oman	−21.7	18.1	97.2	−18.3	15.9	97.9
Cameroon	51.1	61.3	17.3	30.0	35.1	26.9	Switzerland	−19.1	14.4	6.3	−10.6	11.2	8.6
Congo, Rep. of the	49.3	51.2	18.9	27.4	39.4	26.2	Singapore	−15.2	14.3	16.3	6.1	14.7	13.5
Kenya	48.6	55.1	8.5	59.7	61.3	10.2	Saudi Arabia	−12.7	1.8	56.7	−18.4	2.2	57.2
Somalia	45.5	46.4	19.3	34.4	36.2	25.3	United States	−11.2	0.6	14.3	−10.8	0.8	14.5
Uganda	44.3	50.7	8.0	57.7	59.2	13.2	Netherlands	−8.8	13.2	5.4	−6.7	14.1	4.9
Lebanon	42.1	60.4	14.2	57.4	70.7	14.6	Burkina Faso	−8.7	22.0	76.2	11.4	20.7	84.9
Eritrea	41.8	42.4	28.3	46.8	48.4	35.9	Cote d'Ivoire	−7.4	11.2	45.5	−21.8	4.8	19.6
Bosnia and Herzegovina	41.6	45.0	48.2	29.7	34.3	24.2	Libya	−6.5	9.2	18.3	−9.5	12.8	12.5
Mongolia	37.2	42.5	81.2	22.3	25.9	9.8	Sweden	−6.2	5.5	2.4	−3.5	4.8	3.5
Macedonia	36.5	39.0	18.1	30.9	34.9	5.2	New Zealand	−5.4	30.1	1.1	−14.4	24.8	3.1
Congo, Dem. Rep. of the	35.6	46.3	12.4	45.8	55.4	18.9	Norway	−3.0	7.1	2.4	0.6	8.2	1.7
Rwanda	35.1	45.2	31.1	−68.8	70.7	38.0	Spain	−2.6	4.2	13.0	−1.5	3.3	8.5
Sri Lanka	33.5	33.9	13.5	33.0	35.7	30.4	Paraguay	−2.4	7.1	37.0	−2.3	4.7	5.9
Nicaragua	33.3	35.7	12.5	31.3	32.5	7.3	Moldova	−2.2	24.9	61.3	−6.4	8.4	69.1
Nigeria	32.2	37.8	6.4	13.1	15.8	12.3	Belgium	−2.1	7.2	7.2	−0.8	5.9	4.4
El Salvador	32.2	33.1	2.8	33.9	34.6	3.0	Russia	−1.8	10.4	73.7	1.0	5.4	75.5
Malawi	32.0	36.8	43.6	41.9	50.8	30.8	Latvia	−1.7	23.8	43.8	−32.1	31.8	66.9
Togo	31.9	45.6	37.9	19.9	33.9	45.6	Japan	0.4	1.9	7.7	0.9	1.8	4.5
Cuba	31.2	31.5	2.7	32.8	32.9	2.5	Turkey	0.5	7.2	3.9	6.6	12.2	2.8

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.

Table 5 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+). Eight Sub-Saharan African countries belong to the top-25. Other remarkable cases include Jamaica (84.6%), Haiti (80.9%), Laos (45.5%), and Afghanistan (44.2%). Seven other countries that lose more than 30% of their college educated labor force are Bosnia and Herzegovina, Lebanon, Yemen, Macedonia, Sri-Lanka, El Salvador, and Nicaragua. Among the main net receivers, we find many high-income OECD and oil-producing countries but also countries such as Kazakhstan, Paraguay, and Cote d'Ivoire, where relatively few natives have college education.

(b) *Female high-skilled migration*

The migration of highly skilled women is a 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 example Dollar & Gatti, 1999; Klasen, 2000; Knowles, Lorgelly, & Owen, 2002; Coulombe & Tremblay, 2006; Blackden, Canagarajah, Klasen, & Lawson, 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 subsequently lower income levels. Conversely, societies that experience a net female skill gain may experience more favorable growth rates. This issue becomes ever more relevant if developing countries devote significant resources to the education of women in key skill areas to close the gaps with men; but retention is necessary to bear the fruits of these efforts.

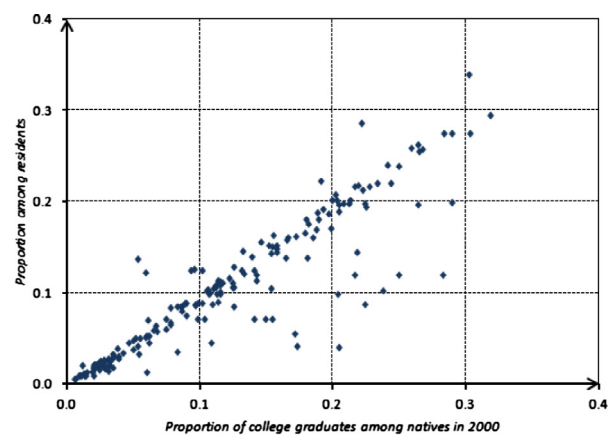
Figure 4 graphically illustrates the impact of our introducing non-OECD destinations into our analysis of female high-skilled emigration rates. In Figure 4.a, we compare the high-skilled emigration of men (*x*-axis) and women (y-axis). Most observations (136 out of 190) lie above the 45° line, indicating that the brain drain is more pronounced in the case of females (as when compared to males). On average, the brain drain for females is 15% higher than for males, as illustrated on Figure 4a by the linear trend estimated for the whole sample.[20] 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 Figures 4b and 3b. Nevertheless, the ratio of non-OECD to total female gross emigration rates in 2000 exceeds 0.50 in 33 countries (as opposed to the 36 cases taking men and women together). Similarly to Figures 3c, 4c 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 4c nevertheless highlights the fact that wealthier countries gain, relative to poorer countries, since they are more successful in attracting higher numbers of college-educated females.

Table 6 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 5, although the magnitude of the net losses are broadly larger for the most affected countries. New entries in the left panel include Mongolia and several African countries, namely, Cameroon, Congo, the Democratic Republic of Congo, Rwanda, Nigeria, Malawi, and Togo; meaning that for these countries the magnitude of high-skilled emigration rates are particularly skewed in favor of women. The only new entries in the right panel in Table 6 (as when compared to Table 5), include Burkina Faso, Norway, Moldova, Japan, and Turkey, meaning that these destinations are particularly attractive to college-educated female migrants relative to their natural female population of college graduates.

(c) *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

a. Women and men together



b. Women only

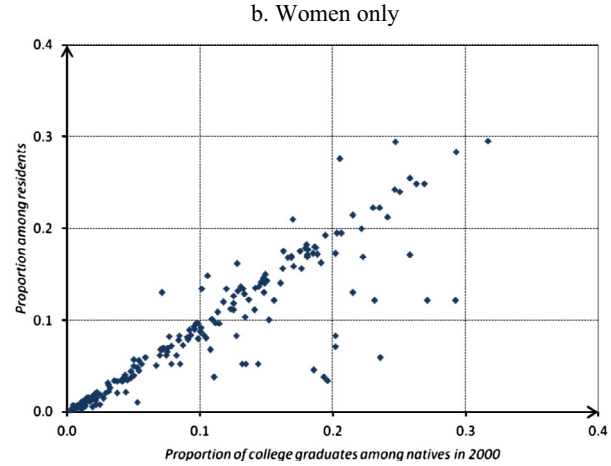


Figure 5. *Human capital among natives and residents in 2000. Note: Human capital is measured by the proportion of college graduates in the population aged 25 and more.*

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 human capital for both the resident and the natural population.

In line with our earlier expression for the average skill levels of naturals and residents, Figure 5a 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° 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 countries 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 41 cases with negative net high-skilled emigration rates, but globally migration only increases human capital in 26 of them. In the remaining 23 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, Gabon, and Russia, but also wealthier countries such as Belgium, France, Luxembourg, the Netherlands, New Zealand, Norway, and the United States among others. Similarly, Figure 5b 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 39 cases with negative net high-skilled emigration rates, but global migration only increases the human capital of females in 26 of them.

6. CONCLUSION

This paper is the first to conduct a comprehensive examination of global human capital mobility, an analysis that rests upon three key contributions, (1) a significant collection of primary data sources, (2) an innovative estimation procedure used to impute data where they are otherwise missing, and (3) a reformulation of existing high-skilled migration measures, which in turn form the basis of our global analysis. Broadening our analysis from the more orthodox approach of focusing solely upon OECD destination countries, yields many important insights that have previously been overlooked. Perhaps most pertinently, migration to non-OECD countries accounts for 20% of all high-skilled migration and these movements comprise relatively large numbers of individuals from low income and least-developed countries in many regions of the world. In the wake of the recent global financial crisis and the shifting balance of power in the global economy, no doubt these migratory patterns will become more pronounced in the years to come.

The database in this paper allows us to paint a picture of human capital mobility around the globe and perform other interesting empirical exercises. We hope our analysis and the data we provide will pave the way for further analytical and empirical work. It is important for those wishing to use the data to familiarize themselves with the methodology we have followed (see Section 3), the adjustments that we have subsequently made (see Section 2) and the detailed list of data sources as described in Appendix A.1. Among the primary data sources, users should have confidence in those that are not marked with an asterisk in Appendix A.1. The bilateral (cell-level) data for these destination countries can be used for empirical analysis with comfort as the adjustments have been minimal. The data for countries marked with an asterisk in Appendix A.1, however, should be treated with more caution, given our preceding discussions. Disaggregated data for this second set of countries can be used to calibrate macroeconomic models for example, but should not be used for econometric analyses. Finally, most attention should be paid to the data for the imputed cells. Aggregating bilateral corridors by origin, destination or region should not pose problems. However, it is clearly not appropriate to use those imputed bilateral data in a gravity model since they are estimated using a gravity model of migration in the first place. Relatedly, and as it is always the case during empirical work, users need take seriously the issue of measurement error. Although our primary raw data comprise the majority of worldwide migrant stocks, imprecision no doubt exists for those cells for which we impute data and is likely to be more severe for smaller corridors. Such imprecision needs to be taken into account when these imputed numbers are included directly in an estimation.

NOTES

1. In the absence of immigration and emigration flow data by skill level, the best countries are able to do in terms of assessing their net human capital situation and thus the effectiveness of their policies, is to compare the total level of human capital at a single point in time (i.e., at the time of census) with the total number of domestic nationals abroad. To be able to do this accurately, bilateral data are required for all potential destinations worldwide.

2. See http://epp.eurostat.ec.europa.eu/portal/page/portal/population/data/database.

3. 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.

4. Note that DLM disaggregated low-skill migrants into 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.

5. http://www.oecd.org/migration/databaseonimmigrantsinoecdandnon-oecdcountriesdioc-e.htm.

6. These regional aggregates recorded in many destination countries' data are usually small.

7. This is the case of Burkina Faso 1996, Colombia 2005, Costa Rica 1984, Israel 1983, Israel 1995, Laos 1995, Malta 1995, Malta 2005, Peru 2007, and Uruguay 2006.

8. In other words, given migrant inflow and outflow data disaggregated by skill level are not available for the vast majority of countries globally, the best available (static) estimates, (thus ignoring dynamic, brain gain effects for example), of the winners of losers of the global battle for talent can be made by comparing the total high-skilled stock in a particular country with the total number of skilled emigrants from that country for which global data on potentially all destinations are required. Even if our imputation methods are imperfect, a sensible allocation of these migrants is still superior for informing one as to the overall global situation, in comparison with the total absence of such data.

9. This IIA (Independence of Irrelevant Alternatives) property has been challenged in a few recent studies (e.g., Bertoli & Fernandez-Huertas Moraga, 2013), which recommends controlling for Multilateral Resistance to Migration. Implementing their corrections is not possible in a cross-sectional setting like ours.

10. Table 8 in the Appendix describes the data sources as well as the way we construct and measure these explanatory variables that influence migrant stocks.

11. A second issue is that the sample of countries for which data are missing is likely a non-random sample. Addressing this second issue is beyond the scope of the current work.

12. Since data for $\lambda_{g,t}^{k}$ need to be available for all 190 destination countries in order for them to be included in our model, the potential exists for some explanatory variables to have been omitted due to data not being available.

13. Youth (for each gender group) is defined as the difference between total migrant stock from OPSW minus low-skilled and high-skilled migrant stock.

14. Similarly, in 1990, we have zero values for 43.2% of the 11,529 observations (190 − 1 destination × 61 origin countries) in the aggregate matrix. The ratio is 46.9% for low-skilled males, 49.0% for high-skilled males, 47.2% for low-skilled females, and 50.9% for high-skilled females.

15. The results are similar for other types and the year 1990.

16. Data on selected non-OECD destination countries were included in Docquier and Rapoport (2012) and in the latest version of the OECD database (DIOC-E).

17. Population data by age and gender are provided by the United Nations Population Division and can be found at http://esa.un.org/unpd/wpp/Excel-Data/population.htm.

18. See https://www.cia.gov/library/publications/the-world-factbook/index.html.

19. We use the gaussian kernel density estimator implemented in Stata.

20. Focusing on OECD destination countries, the gap increases to 18% (see Docquier *et al.*, 2009).

REFERENCES

Anderson, J. (2011). The gravity model. *Annual Review of Economics, 3*(1), 133–160.

Artuc, E. (2013). *PPML estimation of dynamic discrete choice models with aggregate shocks*. World Bank Policy Research working paper No. 6480, World Bank: Washington, DC.

Artuc, E., Chaudhuri, S., & McLaren, J. (2010). Trade shocks and labor adjustment: a structural empirical approach. *American Economic Review, 100*(3), 1008–1045.

Barro, R. J., & Lee, J.-W. (2001). International data on educational attainment: updates and implications. *Oxford Economic Papers, 53*(3), 541–563.

Beine, M., Docquier, F., & Ozden, C. (2011). Diasporas. *Journal of Development Economics, 95*(1), 30–41.

Beine, M., & Salomone, S. (2013). Network effects in international migration: education versus gender. *Scandinavian Journal of Economics, 115*(2), 354–380.

Bertoli, S., & Fernandez-Huertas Moraga, J. (2013). Multilateral resistance to migration. *Journal of Development Economics, 102*(C), 79–100.

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. Helsinki: UNUWIDER.

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

Clemens, M. A., & Pritchett, L. (2008). Income per natural: measuring development for people rather than places. *Population and Development Review, 34*(3), 395–434.

Cohen, D., & Soto, M. (2007). Growth and human capital: good data, good results. *Journal of Economic Growth, 12*(1), 51–76.

Coulombe, S., & Tremblay, J.-F. (2006). Literacy and Growth. *B.E. Journal of Macroeconomics, 6*(2), 1–34.

Docquier, F., & Marfouk, A. (2004). *Measuring the international mobility of skilled workers – Release 1.0*. Policy Research working paper No. 3382. Washington, DC: World Bank.

Docquier, F., Lowell, B. L., & Marfouk, A. (2009). A gendered assessment of highly skilled emigration. *Population and Development Review, 35*(2), 297–321.

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

Docquier, F., & Rapoport, H. (2012). Globalization, brain drain and development. *Journal of Economic Literature, 50*(3), 681–730.

Dollar, D., & Gatti, R. (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., & de la Fuente, A. (2006). Human capital in growth regressions: How much difference does data quality make?. *Journal of the European Economic Association, 4*(1), 1–36.

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

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

Dumont, J. C., Spielvogel, G., & Widmaier, S. (2010). *International migrants in developed, emerging and developing countries: an extended profile*. OECD: Social, Employment and Migration Working Paper 114. Paris.

Grogger, J., & Hanson, G. H. (2011). Income maximisation and the selection and sorting of international migrants. *Journal of Development Economics, 95*(1), 42–57.

Head, K., Mayer, Th., & Ries, J. (2010). The erosion of colonial trade linkages after independence. *Journal of International Economics, 81*(1), 1–14.

Klasen, S. (2000). *Does gender inequality reduce growth and development? Evidence from cross-country regressions*. Collaborative Research Center 386, Discussion Paper 212.

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

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

Özden, C., Parsons, C. R., Schiff, M., & Walmsley, T. (2011). Where on earth is everybody? The evolution of global bilateral migration 1960–2000?. *World Bank Economic Review, 25*(1), 12–56.

Parsons, C. R., Skeldon, R., Walmsley, T. L., & Winters, L. A. (2007). *Quantifying international migration: a database of bilateral migrant stocks*. Policy Research Working Paper No. 4165. Washington, DC: World Bank.

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

APPENDIX A

A.1 *Data sources*

The statistical appendix in DLM (2009, p. 317) describes the data sources for the 30 OECD countries in their sample. Table 7 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.

A.2 *Explanatory variables*

Table 8 describes the data sources for the explanatory variables used in the regression of Tables 2a and 2b.

Table 7. *Migration data for non-OECD destinations (1990–2000 census rounds)*

Country	Source	1990 Round		2000 Round	
		Total	High-skilled	Total	High-skilled
New OECD member states					
Chile (1992–2002)	IPUMS International[b]	15,980	1,930	25,040	4,080
Estonia (1989–2000)	Statistics Estonia	402,958*	113,181*	233,112	72,594
Israel (1983–95)	IPUMS Int'l[b]and Israel CBS	1,178,590	159,800	1,510,067	511,562
Slovenia (1991–2002)	Statistical Office Slovenia	153,953*	16,128*	152,890	17,819
Non-OECD countries					
Argentina (1991–2001)	IPUMS International[b]	54,743	1,512	680,583	60,056
Armenia (n.a.–2001)	DIOC-E database (OECD)[d]	–	–	240,839	55,081
Bahrain (1990–2000)	Labor Force Survey[c]	115,735*	18,295*	153,544*	31,876*
Belarus (1991–99)	IPUMS International[b]	50,931	10,392	946,933	248,826
Benin (n.a.–2002)	DIOC-E database (OECD)[d]	–	–	129,015	1,447
Bolivia (n.a.–2001)	IPUMS International[b]	–	–	45,200	14,560
Brazil (1991–2000)	IPUMS International[b]	341,985	67,229	298,257	67,356
Bulgaria (1991–2001)	National Statistical Institute	16,388*	4,772*	76,951	26,362
Burkina Faso (n.a.–2006)	DIOC-E database (OECD)[d]	–	–	189,188	15,680
China – Hong Kong (n.a.–2000)	Census and Statistics Dep.	–	–	1,854,892	279,965
Colombia (1993–2005)	IPUMS International[b]	41,100	3,400	48,280	5,718
Costa Rica (1984–2000)	I.N. Estadistica y Censos	235,652	29,927	175,454	29,273
Cote d'Ivoire (1987–99)	Institut National de Stat.	3,262,289	30,020	3,906,629	35,916
Croatia (1991–2001)	Central Bureau of Statistics	399,679	58,040	498,153	68,794
Cuba (n.a.–2002)	DIOC-E database (OECD)[d]	–	–	8,770	1,780
Cyprus (1991–2001)	Cyprus Statistics	23,157*	8,672*	42,315	17,095
Dominican Rep (n.a.–2002)	United Nations CEPAL[a]	–	–	37,847	17,681
Ecuador (n.a.–2001)	DIOC-E database (OECD)[d]	–	–	69,134	21,495
El Salvador (n.a.–2007)	DIOC-E database (OECD)[d]	–	–	20,910	4,201
Gambia (n.a.–2003)	DIOC-E database (OECD)[d]	–	–	59,199	3,307
Georgia (n.a.–2002)	DIOC-E database (OECD)[d]	–	–	75,773	19,927
Guatemala (n.a.–2002)	DIOC-E database (OECD)[d]	–	–	25,096	7,583
Guinea (n.a.–1996)	IPUMS International[b]	–	–	126,370	4,920
Honduras (n.a.–2001)	United Nations CEPAL[a]	–	–	17,478	5,635
India (n.a.–2000)	DIOC-E database (OECD)[d]	–	–	5,165,258	147,085
Indonesia (n.a.–2000)	DIOC-E database (OECD)[d]	–	–	6,156	4,708
Iraq (n.a.–1997)	IPUMS International[b]	–	–	50,670	8,450
Jamaica (n.a.–2001)	DIOC-E database (OECD)[d]	–	–	7,541	3,487
Kenya (1989–99)	IPUMS International[b]	39,300	2,080	193,820	12,900
Kuwait (1990–2000)	Labor Force Survey[c]	489,735*	74,780*	668,885*	128,738*
Kyrgyztan (n.a.–1999)	IPUMS International[b]	–	–	312,740	46,200
Laos (n.a.–1995)	DIOC-E database (OECD)[d]	–	–	5,558	468
Latvia (1989.-2001)	Latvia Statistics	675,602*	108,305*	401,471	66,019
Lithuania (1991–2001)	Statistics Lithuania	271,824	41,355	203,374	42,417
Macedonia (1994–2002)	State Statistical Office	43,230	6,198	29,947	5,754
Malaysia (n.a.–2000)	IPUMS International[b]	–	–	769,700	39,400

(*continued on next page*)

Table 7 (*continued*)

Country	Source	1990 Round		2000 Round	
		Total	High-skilled	Total	High-skilled
Mali (n.a.–1998)	DIOC-E database (OECD)[d]	–	–	56,549	2,477
Malta (1995–2005)	National Statistics Office	12,613*	5,279*	19,009	8,524
Mauritius (n.a.–2000)	DIOC-E database (OECD)[d]	–	–	11,067	972
Mongolia (n.a.–2000)	IPUMS International[b]	–	–	4,410	1,440
Morocco (n.a.–2004)	Haut Commissariat au Plan	–	–	34,555	15,247
Nepal (n.a.–2001)	DIOC-E database (OECD)[d]	–	–	391,000	17,665
Nicaragua (n.a.–2005)	United Nations CEPAL[a]	–	–	41,903	4,858
Oman (1990–2000)	Labor Force Survey[c]	279,630*	40,093*	411,640*	75,477*
Panama (n.a.–2000)	IPUMS International[b]	–	–	59,290	11,930
Paraguay (n.a.–2002)	United Nations CEPAL[a]	–	–	105,022	18,408
Peru (n.a.–2007)	DIOC-E database (OECD)[d]	–	–	50,626	29,493
Philippines (1990–2000)	IPUMS International[b]	176,364	69,134	208,517	63,433
Qatar (1990–2000)	Labor Force Survey[c]	194,233*	27,183*	247,201*	45,331*
Romania (1992–2002)	IPUMS International[b]	81,397	24,781	76,519	27,408
Russia (n.a.–2002)	DIOC-E database (OECD)[d]	–	–	9,009,859	2,207,429
Rwanda (1991–2002)	IPUMS International[b]	101,652	9,296	124,550	4,210
Saudi Arabia (1990–2000)	Labor Force Survey[c]	2,842,783*	397,989*	3,078,548*	577,867*
Senegal (n.a.–2002)	DIOC-E database (OECD)[d]	–	–	35,285	6,909
Serbia/Montenegro (n.a.–2001)	DIOC-E database (OECD)[d]	–	–	713,596	114,268
Seychelles (n.a.–2000)	DIOC-E database (OECD)[d]	–	–	3,858	728
Singapore (1990–2000)	Statistics Singapore	397,189	30,191	512,515	137,705
South Africa (1996–2001)	Statistics South Africa	635,110	101,876	795,066	174,873
Sri Lanka (n.a.–2001)	DIOC-E database (OECD)[d]	–	–	14,135	1,729
Tanzania (n.a.–2002)	DIOC-E database (OECD)[d]	–	–	161,390	4,185
Thailand (n.a.–2000)	DIOC-E database (OECD)[d]	–	–	158,445	14,081
Trinidad & Tobago (n.a.–2000)	United Nations CEPAL[a]	–	–	28,225	2,004
Uganda (1991–2002)	IPUMS International[b]	274,198	835	189,700	6,620
Un Arab Emirates (1990–2000)	Labor Force Survey[c]	675,549*	98,565*	1,160,658*	213,445*
Uruguay (n.a.–2006)	DIOC-E database (OECD)[d]	–	–	68,062	10,773
Venezuela (1990–2001)	IPUMS International[b]	493,935	18,243	489,636	37,159

Notes. Exact census years are reported between parentheses after the country name, "n.a." indicated that census data are unavailable. In the last four columns, an asterisk (superscript) indicates that the education–gender structure was unavailable in the census. In most cases, we used the structure of the other census year, adjusting for the change in human capital in the resident population. For GCC countries, we used the structure observed in Saudi Arabia (only available for the total immigration stock).

[a] United Nations' Economic Commission for Latin America and the Caribbean (http://www.cepal.org).

[b] See 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 *Labur Force Sample Survey* at http://www.qsa.gov.qa, for Bahrain see *Labor Market Indicators* at http://blmi.lmra.bh, for Oman see Periodic *Labor Force Survey* at http://www.moneoman.gov.om and for Kuwait see *Microdata of the Labor Force Survey* at http://scs.mop.gov.kw.

[d] Please refer to "Dumont, Spielvogel & Widmaier (2010). Les migrants internationaux dans les pays développés, émergents et en développement: élargissement du profil, Questions sociales, emplois et migrations, n. 114."

Table 8. *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

[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.

APPENDIX B. SUPPLEMENTARY DATA

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.worlddev.2014.04.004.