Geographic, Gender and Skill Structure of International Migration

Docquier, Frédéric and Marfouk, Abdeslam and Özden, Caglar and Parsons, Christopher

IRES-UCL, FNRS, IWEPS and Université Libre de Bruxelles, World Bank, Nottingham University

September 2011

Geographic, Gender and Skill Structure of International Migration*

Frédéric Docquier (IRES-UCL, FNRS), Abdeslam Marfouk (IRES-UCL),

Çağlar Özden (World Bank) and Christopher Parsons (Univ Nottingham)

September 2011

Abstract

This paper presents a new bilateral database documenting international migration stocks by gender, education level, origin and destination. We build on existing databases of OECD host countries in 1990 and 2000 and expand their coverage by collecting or estimating migration to all non-OECD destinations. The end result is comprised of comprehensive 195x195 matrices of international migration for 1990 and 2000, distinguishing migrants by gender and education (college educated and the less educated). This unique database allows us to characterize the origin, destination and education levels of about 100 million adult migrants in 2000, and 80 million in 1990. We identify the main sending and receiving countries and the largest migration corridors. South-North corridors are dominant in 2000, exhibiting the largest growth rates and skill intensity. Accounting for migration to high-income non-OECD countries and numerous developing countries drastically improves our understanding of global migration patterns.

JEL codes: F22, J61, O15.

Keywords: International migration, labor mobility, brain drain.

*Thanks are due to the the Economic Research Forum for its financial support (contract #2009-057 on "The efficiency and redistributive effects of international labor mobility: a bilateral approach with special focus on MENA migration"). The first two authors also acknowledge 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. Email addresses: frederic.docquier@uclouvain.be; a.marfouk@skynet.be; cozden@worldbank.org; chrisparsons25@yahoo.co.uk. Tel. +32 10 474 149. Fax: +32 10 473 945.

1 Introduction

International migration is a powerful force that shapes the distribution of human populations across the globe thereby affecting their social, political, and economic structures. Many aspects of migration have been analyzed by demographers, economists, sociologists, and other social scientists. Indeed, specialized academic journals are entirely devoted to its analysis. However, due to the absence of comprehensive and detailed data, our knowledge of the bilateral composition of international migration and its effects is less than desirable. These data shortcomings impede many important avenues of research such as the determinants of various dimensions of migration patterns (such as gender, age and skill composition), reasons behind the emergence and disappearance of important migration corridors, and the analysis of linkages between migration patterns and social and economic development.

Disentangling the volume of migration by country of origin, country of destination, gender and education level is crucial in understanding its demographic, economic, political and sociological consequences. Demographically, international migration is one of the key components governing population dynamics. The construction of bilateral databases is among the key steps to understanding the process of migration, especially in identifying its causes and consequences. Migration influences the dynamics of societies by altering the age, gender, skill compositions in the origin and destination countries.[1] It has important effects on fertility and mortality rates. Migration may induce transfers of cultural, sociological or behavioral norms between countries. [2] Analyzing both the direct and indirect demographic effects, together with their dynamic implications, is only possible with bilateral data disaggregated by education and gender over time.

International migration has historically received scant attention when compared to other aspects of globalization, especially in the economic and political science literatures.[3] Winters (2001), Winters and Walmsley (2005) and Pritchett (2006)

[1] The effect of migration on host-country demographics has been comprehensively studied in the literature. See Goldberg (1960) and Freedman and Slesinger (1961) among others.

[2] For example, Fargues (2007) shows that fertility rates in Morocco, Turkey and Egypt are affected by the rates prevailing in their migrants' host countries. This hypothesis receives empirical support in Beine et al. (2008).

[3] A significant body of literature investigates the economic linkages between migration and other aspects of globalisation. Gould (1994), Head and Ries (1998), Rauch and Trindade (2002), Rauch and Casella (2003), Combes et al. (2005) and Peri and Requena (2009) amongst others, document the extent to which migrants foster bilateral trade flows. A comparable body of literature investigates

have argued that international migration should be seen as the key to unlocking the true gains from globalization. Laws and regulations which restrict migration from the South to the North carry considerable economic welfare costs for both developed and developing countries and serve to compound existing income inequalities and poverty traps. Even in this restrictive policy environment, remittances sent back home by overseas workers, moved to the forefront of the development dialogue. In 2009, remittance flows to developing countries are estimated to have reached some $316 billion (Mohapatra et al. 2010) which dwarves the total development assistance from all OECD countries which only amounted to 132.7 billion in the same year.[4] Obviously, the costs and benefits of international migration depend upon the skill and gender structure of migration patterns. Again, an accurate assessment of those effects requires broad, disaggregated and reliable data.

In his treatise on "Exit, Voice and Loyalty", Hirschman (1970) envisaged migration as a political response to failing institutions and discusses its interplay with protest and patriotism. More recently, the link between migration and the quality of institutions has received attention. Spilimbergo (2009), for example, shows that foreign-trained individuals promote democracy at home, so long as foreign education is acquired in democratic countries. Exploratory regressions in Li and McHale (2006) show that the brain drain has a positive effect on domestic "political" institutions, but a negative effect on home country "economic" institutions. Docquier et al. (2010) show that openness to migration and human capital have a positive impact on institutions (as measured by standard democracy and economic freedom indices) using dynamic panel regressions. Detailed data used in these studies are only available for a limited number of country pairs. Again, additional global data disaggregated by education level are needed to better understand these mechanisms.

This paper addresses these shortcomings by presenting, for the first time, a global analysis of bilateral migration patterns by gender and education levels. In particular, (compared to previous analyses), we account for migration to all developing country destinations by introducing new census data and utilizing appropriate estimation methods. The database constructed and presented in this paper complements and extends other data sources in the literature. The most popular, the United Nations'

the extent to which migrants beget FDI (see Kugler and Rapoport, 2006, Javorcik et al., 2010). Importantly, these links vary depending upon the skill or educational attainment of migrants.

[4]See http://www.oecd.org/document/11/0,3343,en_2649_34447_44981579_1_1_1_1,00.html

International Migrant Stock database provides time series data on the stock of immigrants, by country, but has no bilateral dimension. This aspect is examined in the Eurostat[5] database and in Ozden et al. (2011), referred to as OPSW henceforth). The Eurostat database provides incomplete data on the size of migration flows, by age, gender and country of citizenship, for migrants in the EU member states only. More broadly, OPSW construct five 226x226 matrices of origin-destination stocks, which correspond to the last five completed census rounds, thereby extending the work of Parsons et al. (2007). However, while OPSW significantly broadens the time, gender and geographical coverage of the available data, different skills or education levels are not distinguished.

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

The existing databases of bilateral migrant stocks disaggregated by education level only capture the size and structure of migration to the whole set, or to a subset, of OECD destinations. This is an important limitation, since OECD nations only host 47 percent of international migrants according to the United Nations. Countries such as Russia, Ivory Coast, Pakistan and India also attract large numbers of migrants, including many from neighboring countries. In terms of high-skill migration, countries such as South Africa, the member states of the Gulf Cooperation Council (referred to as GCC henceforth) and some East Asian countries (e.g. Singapore or Hong Kong) are among the most important non-OECD destinations. Omitting these destinations from a database means that an important piece of the global puzzle goes missing, thereby limiting our understanding of the full nature of international human capital mobility.

The purpose of this paper is to bridge this gap in the literature by providing the first comprehensive database of migration by skill level and gender. We collect or estimate data from non-OECD destinations to expand the coverage of existing

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

studies. We start from DLM and increase the number of host countries, adding 46 non-OECD destinations in 2000 and 31 in 1990. Utilizing these observations, we then estimate the size and structure of immigration in the rest of the world using gravity regression models, which account for the frequency as well as the magnitude of migration flows.

When the recently collected and estimated data are all combined, the final output is a unique database characterizing the origin, destination, education level and gender of about 100 million adult migrants in 2000, and 80 million in 1990. We provide comprehensive 195x195 matrices of international migration for both of these years and distinguish migrants with college education (i.e. post-secondary education) from those with lower degrees as well as males from females. Our database provides accurate measures of all low-skill and high-skill migration flows internationally and it is the first database to characterize true "South-South" and "North-South" emigration patterns by these characteristics.

The remainder of this paper is organized as follows. Section 2 describes the methodology and data sources used. Some key findings are then presented in Section 3. Here we identify the most important global patterns. Finally, Section 4 concludes.

2 Sources and Method

The contribution of this paper is the construction of comprehensive bilateral migration matrices by education level and by gender for 1990 and 2000. As discussed above, these matrices can be used to evaluate human capital mobility across the globe over time. They cover 195 origin countries and focus on stocks of migrants aged 25 and above. This cutoff is chosen due to data avaliability from national censuses as well as our desire to focus on labor market dynamics by leaving out students and children. The data are disaggregated by gender and two separate education levels so that, in each decade, we have migrant stocks of high-skilled males, low-skilled males, high-skilled females and low-skilled females for every bilateral corridor.

This section describes the data sources and the methodology used to construct the migration matrices. As explained below in sub-section 2.1, we start with the existing database of DLM, which documents migrant stocks disaggregated by education levels from 195 origin countries to 30 OECD destination countries. We use the same definition as in DLM and add 46 non-OECD destinations for 2000 and 31 countries

for 1990 for which parallel data could be found from original sources, mainly national statistical agencies. Next, we use the primary data from these 195x76 and 195x61 matrices to predict the size and educational structure of migration stocks for the remaining 119 destination countries in 2000 and 134 countries in 1990 as described in sub-section 2.3.

Before delving into the details of the empirical exercise and the analysis of our data, we present summary statistics of the numbers of high-skilled migrants in the database in Table 1, while distinguishing between raw data made available by national statistical offices and estimated/imputed data. For each year, the migrant stock in the 30 OECD countries is shown in the first column. There are 57.4 million migrants above age 25 in 2000, of which 20.3 million (35.4 percent) are highly educated, and 29.2 million (50.9 percent) are women. For 1990, we identify 40.8 million migrants including 30.0 percent highly educated and 50.6 percent women.[6]

The second column shows the data obtained from primary sources for non-OECD countries. In the 46 destination countries for which we obtained data for 2000, there are 20.2 migrants, of which 3.1 million (15.1 percent) are highly educated and 8.4 million (41.3 percent) are female. For 1990, we identify 14.0 million migrants, including 11.3 percent highly educated and 38.6 percent women. Compared to OECD destinations, the shares of both the low-skilled and female migrants in non-OECD countries are lower.

Finally, for completeness, the third row presents the results of our estimations which are discussed in detail below. For 2000, we predict 22.9 million migrants in the 119 destination countries for which actual data are not available. Of these, only 2.8 million (12.2 percent) are predicted to be high-skilled, and 11.5 million (47.0 percent) are women.

Overall, the migration matrices identify 100.5 million migrants (age 25+) in 2000 which represents about 56.7 percent of the 177.4 million migrants (age 0+) recorded in the United Nations database and 62.8 percent of the 160.1 million migrants (again age 0+) recorded in OPSW for those 195 countries that appear in our matrices.[7] 26.2 million of this migrant stock has college education and 49.0 million are women.

[6] The small differences with DLM are due to the exclusion of migrants for which data by education levels are available, but information on their country of origin is missing.

[7] 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.

For 1990, we identify 80.2 million migrants (aged 25+), including 16.4 million high-skilled migrants and 38 million women. Our data indicate that the overall migrant stock increased by 25.3 percent between 1990 and 2000, while the stock of high-skilled migrants increased by 60.1 percent. As a result, the share of high-skilled in the overall migrant stock increased from 20.4 percent to 26.1 percent. The share of women increased from 47.3 percent to 48.8 percent, a result mainly driven by the increased feminization of migration to non-OECD countries.

[INSERT TABLE 1 AROUND HERE]

2.1 Migration to OECD countries

DLM (2009) database is based on a collection of census and register data by country of birth, education level and gender in the 30 OECD countries in 1990 and 2000. This original set of OECD destinations does not include the three member states which joined the OECD in 2010 (Chile, Israel and Slovenia) and future members (Estonia and Russia). DLM counts stocks of migrants living in the destination country at the time of census, not flows observed between two census dates. For obvious reasons of homogeneity and comparability, the methodological choices made in DLM guide our extensions. The four main principles are the following:

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

- Migration is generally defined on the basis of country of birth rather than citizenship. This definition is time invariant (contrary to the concept of citizenship, which changes with naturalization) and independent of the variation in laws regarding citizenship within and across countries.

- Only adult migrants aged 25 and above are recorded. This excludes students who temporarily move abroad to complete their education or children who migrate with their parents. This is a better measure of the labor market and economic impact of migration.

- Besides the gender dimension, two separate levels of education are specified. High-skill migrants include those with college or post-secondary education, i.e. any degree above upper-secondary level. Low-skill migrants are those with upper-secondary education and less.[8]

Table A1 in the appendix of DLM (p. 317) describes the data sources for the OECD. As shown in Table 1, the OECD data allow us to characterize the education level, origin and destination of about 57.4 million migrants in 2000 and 40.8 million migrants in 1990. About 16.9 million of the 20.3 million high-skill migrants in the OECD countries are concentrated in only 5 destination countries: the U.S. (10.3 million), Canada (2.7 million), Australia (1.6 million), United Kingdom (1.2 million) and Germany (1.2 million).

2.2 Migration to selected non-OECD countries

In this paper, we collect immigration data for 46 non-OECD countries in 2000 and 31 countries in 1990 following the same principles and definitions in DLM. This non-OECD group includes the following countries where a superscript * indicates that data are missing for 1990 and a superscript ° indicates that the country has joined the OECD since 2000:

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

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

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

Adding 46 destination countries increases the migrant stock by 20.2 million in 2000, including 3.1 high-skilled migrants and 8.4 million women, as shown in Table 1. The average ratio of high-skilled is 15.1 percent and the share of women is 41.3 percent, much below the ratios observed in OECD destination countries in both of these dimensions (35.4 percent and 50.9 percent, respectively). Those ratios vary considerably across countries and this heterogeneity is explored in more detail in section 3.[9] Six of these 46 additional destination countries are home to more than one million foreign-born adults in 2000. These are Ivory Coast (3.9 million), Saudi Arabia (3.1 million), Hong Kong (1.9 million), Israel (1.5 million), the United Arab Emirates (1.2 million) and Malaysia (1.0 million).

[9] In addition, we have 715 thousand migrants (of which 2467 thousand are highly skilled) for whom data by country of origin are missing. These migrants are not included in the numbers reported in Table 1.

2.3 Migration to other non-OECD countries

Even with the addition of data on 46 destination countries in 2000 (and 31 countries 1990) to the already existing data from 30 OECD countries, our bilateral migration matrices are far from complete. The goal of this section is to predict bilateral migrant stock data by both gender and education for the missing 119 non-OECD countries in 2000 and the 134 countries in 1990. We first estimate the determinants of the bilateral migrant stocks (disaggregated by gender and education level) for which we have data. Then we use the parameter estimates from these regressions to predict the gender and education specific bilateral migrant stocks for the missing cells of our matrices. This section concludes with analysis of the accuracies of our predictions.

While analyzing the effect of diasporas on future migration flows, Beine et al. (2010) derive a gravity-type equation from an individual utility-maximization model with wage differentials. In the model, individuals with different human capital (education) levels choose between alternative destinations and staying at home after observing their individual random shocks. Each country pair (or corridor) is also characterized by pair-specific migration costs and barriers such as physical distance, linguistic overlap and political linkages. This model leads to a gravity-type equation where gender and education specific migration levels are expressed as a function of various bilateral variables as well as destination and origin country specific push and pull variables. We use the following specification in our estimation of the determinants of bilateral migration patterns:

$$M_{ijt}^{g} = \alpha_{0t}^{g} + \alpha_{1t}^{g} d_{ijt}^{g} + \alpha_{2t}^{g} A_{jt}^{g} + \gamma_{it}^{g} + \epsilon_{ijt}^{g} \tag{1}$$

The variables are defined as following:[10] The main dependent variable M_{ijt}^{g} is the bilateral migrant stock in the gender-education group g from country i to country j in the relevant year t - either 1990 or 2000. We have two years and four separate gender-education pairs - high-skilled men, high-skilled women, low-skilled men and low-skilled women. We run eight separate regressions to estimate year- and group-specific parameters and fixed effects.

The explanatory variables include a set of bilateral geographic distance and linkage (linguistic, political, cultural) variables denoted by d_{ijt}^{g}. Among these are geographic

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

contiguity, physical distance, colonial relationships and linguistic overlap (see Clair et al., 2004). We also include the total number of migrants by gender from country i in country j at time t based on the bilateral migration database presented in OPSW.[11]

All origin country characteristics (such as economic, political and social push factors) are captured via fixed effects presented through the vector γ_{it}^{g}. Ideally, we would also like to include fixed effects to capture all destination country specific pull factors and, thus, increase the predictive power of the bilateral variables and the overall estimation. Unfortunately, the inclusion of destination fixed effects would prevent us from predicting the values of the missing cells in the migration matrices since the coefficients of these dummy variables are of no use in making predictions. Instead, we use a set of ten regional dummies (South Asia, Latin America etc.) as well as various destination specific variables that influence migration patterns, which are denoted by A_{jt}^{g}. These include whether the people in the destination country speak English, the size of the total labor force in the destination country (in log), GDP per capita (in log), total fertility rate (in log) in the destination country, the ratio of the number of skilled to the total in the destination country and the labor force participation rate in the host nation. A number of dummy variables are also included that capture whether a destination country belongs to the GCC, whether military service is compulsory in the destination and whether polygamy is legally practiced. We believe that together with the origin country and destination region dummy variables, these capture the most important determinants of international migration. It is worth noting that all of the destination variables, A_{jt}, must be available for all 195 destinations in order to predict the missing migration numbers.

2.3.1 Econometric issues

The estimation of (1) entails various econometric issues that would lead OLS to generate inconsistent estimates. The most important of these is the presence of a large number of zero or undefined observations for the dependent variables (gender and education specific bilateral migrant stocks) in both 1990 and 2000. Zero observations

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

appear in large numbers in many other 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, from the Central African Republic to Peru, due to high geographic, cultural or economic barriers. Furthermore, even if there were minimal migration, censuses or other survey instruments might not capture it due to limited sampling. As a result, we have zero values for about 48.5 percent of the 14,820 observations (195 destination x 76 origin countries) in the aggregate migration matrix for 2000. The ratio of zero observations is 52.6 percent for low-skilled males, 52.9 percent for high-skilled males, 52.8 percent for low-skilled females and 54.0 percent for high-skilled females.[12]

If OLS were used in the estimation of (1) with the size of migrant stocks as the dependent variable, the estimates are likely to be inconsistent. One alternative is to use the natural log of the migrant stocks. Zero observations are dropped from the sample in such specifications since the natural log of zero is undefined. In that case, the results are likely to be biased and the impacts of the explanatory variables are likely to be underestimated due to the exclusion of low value observations from the sample. Alternatively, we can add one to a migrant stock and then take the natural log. This also leads to heteroskedasticity in the estimation, however, since the log of one is zero, which again leads to an excessive number of zero observations in the dependent variable.

The most appropriate solution to this problem is to use Poisson regression models that rely on pseudo-maximum likelihood estimates. Santos Silva and Tenreyro (2006) in their influential paper show that the log linearization of gravity models leads to inconsistent estimates of coefficients of the bilateral explanatory variables such as distance. As a result, we implement Poisson regressions for the model explaining the size of the gender and education specific migrant stock in (1). Importantly, however, all the Poisson models are estimated with robust standard errors to mitigate a further econometric complication. This relates to the fact that Poisson maximum likelihood estimation yields consistent point estimates even when the count is not strictly Poisson distributed i.e. in cases of over-dispersion. Importantly, in such

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

circumstances, the estimated standard errors will be significantly smaller than if the count *was* strictly Poisson. This occurs when the conditional variance is greater than the conditional mean i.e. when the assumption of equal dispersion is violated. Observations are weighted by the log of the aggregate migrant stock and use of robust standard errors addresses this problem.

2.3.2 Results

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

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

As far as education is concerned, we find evidence of migrant selection. Countries with a higher proportion of highly skilled workers attract higher proportions of highly skilled migrants. The oil-rich GCC (Gulf Cooperation Council) countries also attract larger numbers of highly skilled as do countries with larger labor forces, although this also deters less educated females. The existing literature (see Grogger and Hanson, 2008, and Beine et al., 2010), shows that the educational composition of migrants from more distant countries is biased towards the more skilled. This would be in accordance with our findings from the contiguity variable, but this is not confirmed by our distance variable.

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

[INSERT TABLE 2 AROUND HERE]

2.3.3 Accuracy of in-sample predictions

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

Table 3 provides summary statistics of our in-sample predictions for the four groups of migrants in the year 2000.[13] In each section of Table 3, the first column provides the number of observations. The second column *'Corr'* provides the correlation between actual and predicted values. The third column *'Migs'*, gives the total number of actual migrants (in millions) while the fourth gives the migrants as a percentage of the relevant gender and education specific sample. The final two columns provide the means and the standard deviations of the logarithm of the ratio of our predicted values to the actual numbers. These are the same things as the differences of the logs which arguably provide the best summary statistics of exactly

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

how accurate our predictions are.[14] These statistics are compiled for different sizes of migration corridors. The first row is for all observations; the second, third and fourth rows are when the predicted values and the actual values are strictly less than 250, when the predicted and actual values are greater than 250 and when both are greater than 5000. The fifth and sixth rows are for OECD and non-OECD destinations when the predicted and actual values are both greater than 250.

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

The correlations between the predicted and the actual migrant stocks present only part of the picture. The next exercise analyzes the log of the ratio of the predicted stocks to the actual stocks of migrants. If our predictions were flawless, all of these ratios would be equal to 0. In Table 3, columns *'Mean LR'* and *'Sd LR'* give the unweighted means and standard deviations of these log ratios by corridor size and higher values indicate lower precision. The most striking numbers are in the first and second rows. The mean of the log ratios range from 0.18 (for low-skilled males) to 0.32 (high skilled females), implying significant deviations of the predicted values from the actual ones. However, when we focus upon larger corridors, the summary statistics of

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

these log ratios improve dramatically. For example, for corridors above 250 (row 3), the mean of the log ratio ranges between 0.05 to 0.10. Since these groups comprise over 99 percent of our in-sample migration stock, we conclude that our estimates imply relatively small global biases overall.

[INSERT TABLE 3 AROUND HERE]

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

Each figure presents graphically the distribution of the (log) ratios on the real line. A distribution closer to 0 and with minimal variation implies more precise estimation.[15] We present the results for all corridors, less than 250, greater than 250 and greater than 5000 (first four rows in Table 3). These figures highlight the large inherent idiosyncratic factors that exist in the estimation of small migration corridors. As clearly seen in the comparison of Figures 1b, 1c and 1d, the distribution of small corridors cover a much wider range indicating a larger standard deviation. On the other hand, as the corridors get bigger, the distribution of the log ratio becomes more concentrated around zero and symmetrical. This once again shows the increasing precision in the prediction of larger corridors.

[INSERT FIGURE 1 AROUND HERE]

3 Main Patterns in the Database

The previous section describes how we use recorded migration stocks (when census and register data are available) for 76 destination countries in 2000 to predict the data for 119 destinations for which official statistics are missing.[16] We then combine the recorded original data with the predicted numbers to construct the global database which has gender and skill specific migrant stocks for each one of the 195x195 corridors

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

[16] Similarly, we use data from 61 countries in 1990 to predict the missing data for 119 destination countries.

for both 1990 and 2000. Our final correction is to set the migrant stocks to zero if the predicted (gender and skill specific) migrant stock in a given corridor is below 250. As explained above, this is due to the lack of precision in the estimation of small corridors which account for a relatively large number of observations but less than one percent of the overall migrant stock.

Our database complements the United Nations' international migrant stock database (which provides data only by destination) and the bilateral databases developed by Parsons et al. (2007) and OPSW. The main difference is that we document bilateral movements of adult workers (age 25+) by both gender *and* education level on a comprehensive basis. This enables us to identify the main sources and destinations of migrants, characterize the bilateral structure of migration patterns for the world and identify the most important corridors.

In this section, we provide some general statistics and highlight the key features of the data. We divide our analysis into three parts. The first examines the main regional trends. We then identify the main sources and destinations with a specific focus on the human capital intensities of international migration. Finally, we investigate the most important global migration corridors.

3.1 Regional aggregates

Table 4 details the total immigrant and emigrant stocks and their composition in 2000 for key regions of the world. The top portion of Table 4 first divides the world into high-income and developing countries. We, then, distinguish between low income, least developed and small island developing states (SIDS) which have unique migration patterns. The second section of the table divides the world into twelve geographical regions: (1) the United States, (2) Canada, Australia and New Zealand as a single entity which is referred to as CANZ, (3) the oil rich Gulf Cooperation Council (GCC) countries (4) the twenty-seven nations of the European Union, (5) Latin America and the Caribbean, (6) Sub-Saharan Africa, (7) the countries of the Commonwealth of Independent States, (8) India, (9) China, (10) countries in the Middle East and North Africa excluding the GCC, (11) other countries in Asia, (12) and a category for all remaining (small) countries.

High income countries attract the largest absolute number of migrants (67.8 million). 33 percent of these are college educated (as opposed to 11.6 percent in develop-

ing country destination) and 48.7 percent are women (which is almost identical to the level in developing countries). There are 42 million migrants coming from developing countries and living in high-income countries; this number includes 13 million high-skilled migrants and almost 20 million women. The proportion of college educated immigrants declines as the income of destination nations decreases with 52.3 percent of immigrants in CANZ having tertiary education. This pattern is also reflected in the skill ratio among emigrants. High-income countries' emigrant stocks are more highly skilled than their immigrant stocks, although obviously they attract far greater numbers of immigrants than they themselves send abroad, especially in the case of the USA and CANZ. Compared to 1990, the stock of high-skilled immigrants in developed countries increased by 36.1 percent, while the stock of their emigrants increased by only 9.1 percent as of 2000. This shows the increasing tendency of human capital to agglomerate in locations where it is already relatively abundant.

[INSERT TABLE 4 AROUND HERE]

As far as the geographic distributions of immigration and emigration are concerned, the United States remains by far the most important single destination for migrants, particularly for the highly skilled. While the European Union is almost comparable in terms of the overall numbers of immigrants, a far lower proportion are high-skilled. Moreover, far greater numbers of EU nationals emigrate as when compared to their American counterparts. Finally, while all the wealthy regions have net low-skilled migrant stocks (net immigration stocks amount to 7.1 and 5.3 million in CANZ and GCC countries, including 3.6 and 1.0 million college graduates), the European Union actually records a net outflow of 1.5 million high-skilled workers.

Among the countries of the developing world, Asia and Latin America send the highest numbers of highly skilled workers abroad, while Sub-Saharan Africa and Latin America send the largest volumes of low-skilled migrants. In parallel important increases in emigration levels are observed in Latin American and the Caribbean, Asia and Sub-Saharan Africa.

3.2 Main sending and receiving countries

Figure 2 visually demonstrates the relative size and skill structure of international migration in all countries around the world in 2000. Figures 2a and 2b represent

immigrant and emigrant stocks as percentages of the destination and origin country populations, respectively. Darker colors represent a larger proportion of migrants in the relevant population. Unsurprisingly, Figure 2a shows that rich countries (the US, Europe, CANZ, or GCC countries) exhibit larger immigration rates, while Asia, Latin America and sub-Saharan-Africa are at the opposite end of the spectrum. Nonetheless, some developing countries, such as Russia, Gabon, Ivory Coast, Gambia or Ghana, also have remarkably large immigration rates, mainly due to regional flows.

Our database can also be used to assess the size of emigration. Figure 2b shows that emigration rates are lower in high-income countries, although there is considerable emigration from Western Europe and Canada. The main source countries, in relative terms, are small developing countries in Western Africa (in particular, Mali, Mauritania and Burkina Faso), Central America and the Caribbean. These are parallel to the conclusions in OPSW.

A more original feature of our database is that it can be used to characterize the gender and skill composition of immigration and emigration stocks. Although the gender composition is relatively homogenous across regions (with a few exceptions such as GCC destination countries), the skill composition is much more diverse. Figures 2c and 2d represent the proportion of college graduates in the immigrant and emigrant stocks in 2000, respectively. As seen in Figure 2c, North America, Australia and New Zealand have the highest ratio of the educated among the immigrants. Positive selection, to a lesser degree, is also observed in Europe. The skilled immigrant shares are also large in Morocco, Egypt and Papua New Guinea, although these high proportions apply to a small overall immigrant base (as shown on Figure 2a). Finally, Figure 2d depicts positive selection in emigration. The proportion of college educated emigrants is large in all high-income countries and small in Sub-Saharan Africa, a reflection of the skill level of the underlying labor supply. Nevertheless, positive selection is strong in many developing countries such as Chile, Argentina, Venezuela, South Africa and Iran.

[INSERT FIGURE 2 AROUND HERE]

Table 5 completes the picture, presenting the largest countries of immigration and emigration globally. In addition, the gender/skill composition of the migration stocks, and their evolution observed between 1990 and 2000 are included. We first describe immigration patterns (top panel). In absolute terms, the US is by far the largest

immigration destination for both overall (24 million) and for high-skilled migration (10 million). Other high-income countries such as Germany, the United Kingdom, France, Canada and Australia are among the main destinations. Due to quality-selective immigration policies, Canada and Australia belong to the top-3 receiving countries of high-skilled workers. These four English countries - the US, Canada, Australia and the UK - receive 15.8 million of the 26.2 or over 60 percent of the high skilled migrants worldwide.

Lagging far behind the United States, the second largest country of immigration is Russia (4.9 million) with a high-skill ratio of only 18 percent. Hence, Russia falls to the 6th place in the ranking based on the number of high-skilled immigrants. Saudi Arabia belongs in the top-10 in terms of total and high-skilled immigration, yet, compared to other destinations, the share of women is extremely low (at 11 percent). It is worth noting that two sub-Saharan African countries (Ivory Coast and Ghana) have large immigration stocks, with very low proportions of college educated (0.9 and 1.6 percent, respectively). The last four columns list the countries with the highest immigration levels relative to their resident populations. They range from 40 percent in Bahrain to over 80 percent in Macao. This group includes relatively small and rich countries (Macao, Israel, GCC small states) as well as several African countries (Gabon and the Ivory Coast).

As far as emigration is concerned (bottom panel), Mexico is the single largest source with 6.5 million citizens living abroad, predominantly in the US. Among the main origin countries, we find several large developing countries (such as India, China, the Philippines, Russia and Ukraine), developed European countries (United Kingdom, Italy, Germany), and finally, Burkina Faso (which sends many migrants to Ivory Coast). Restricting the analysis to high-skill migration, Mexico falls to 6th place (only 15 percent of Mexican emigrants have college education) and the Philippines jumps to the 3rd place . Korea and Vietnam are among the top-10 exporters of high-skilled labor. In relative terms (see last four columns), the highest emigration rates are observed in small countries. Excluding microstates with less than 100,000 inhabitants, our top-10 is made of small and relatively poor countries located in the neighborhoods of large destination countries. For example, six countries have lost more than 50 percent of their working-age population (Suriname, Burkina Faso, Guyana, Cape Verde, Jamaica and Togo) as of 2000. Proportions of college educated are relatively low in this group, except in Guyana, Jamaica and Barbados. Shares of women are

close to 50 percent, except in Mauritania (23 percent).

[INSERT TABLE 5 AROUND HERE]

3.3 Main corridors

The most valuable feature of this database is its bilateral nature and this section identifies the main migration corridors, and characterizes their gender and skill structures. Table 6 lists the 40 most important corridors which account for 0.1 percent of the 38,025 (195x195) observations in our database but a massive 36.1 percent of the 100.5 million migrants identified in 2000. This reflects a strong concentration in international migration patterns in which a few large corridors account for a disproportionate share of the overall migrant stock. The United States appears 15 times as a destination in this list, with a share of college educated varying between 14.4 percent for Mexicans to 79.4 percent for Indians. Many of the largest corridors are between countries sharing common borders (Mexico to the US, Burkina Faso or Mali to Ivory Coast, Ukraine or Kazakhstan to Russia, China to Hong Kong, Bangladesh to India, Togo or Burkina Faso to Ghana, etc.) where the proportion of college graduates is usually very small. On the other hand, when large corridors involve distant pairs, the skill share is usually high (71.7 percent for the Philippines-the U.S., 79.4 percent for India-the U.S., 39.4 percent for U.K.-Australia, etc.). In line with Grogger and Hanson (2008), Beine et al. (2010) and as confirmed by our regression results from the previous section, these patterns indicate that low-skilled migrants are more sensitive to distance and less responsive to income differentials than the high-skilled. The main exceptions are distant pairs of countries with past colonial links or guest-worker agreements. For example, the proportion of college graduates equals 6.3 percent for Turkey-Germany, and 10.2 percent for Algeria-France. The share of women are relatively close to 50 percent in most large corridors, except for migrants to Saudi Arabia, or those from Togo to Ghana.

The right-hand-side of Table 6 presents the largest 40 corridors based on net flows which are calculated as the change in the total stocks between 1990 and 2000. These corridors comprise 12.1 million migrants which is 57.3 percent of the 21.2 million new migrants identified between 1990-2000. The United States appears 16 times as a destination. There is a sharp increase in immigration of the low-skilled from Mexico (3.7 million) and various other Central American countries (such as

El Salvador, Dominican Republic and Guatemala); and in high-skilled immigration from Asia (India, Vietnam, the Philippines, China, Korea). Increasing migration is observed in corridors involving ex-Soviet block country pairs, such as between Russia and Kazakhstan. Migration to Ghana, United Arab Emirates and Hong Kong is also on the rise.

[INSERT TABLE 6 AROUND HERE]

The bilateral nature of our database enables us to quantify the exchanges between developing countries (henceforth referred to as the South) and high-income countries (referred to as the North). Based on the 2000 migration stocks (left columns in Table 6), our top-40 corridors include 16 South-North (40 percent), 13 South-South (32.5 percent), 11 North-North (27.5 percent), and no North-South corridors. Based on 1990-2000 migration net flows (right columns in Table 6), our top-40 corridors include 25 South-North (62.5 percent), 13 South-South (32.5 percent), only 2 North-North (5 percent) and no North-South corridors. Table 7 generalizes this trend which indicates the increasing importance of South-to-North migration in the 1990s. First, despite the fact that there are many more developing than high-income countries, the stock of South-North migrants (45.4 million) exceeds the South-South (29.4 million) and North-North (22.5 million) stocks. North-South movements are negligible. Second, it is also in South-North corridors that the highest growth rates are observed, +55.6 percent while 10 percent for the other pairs. The lowest growth rates are observed in South-South migration.

[INSERT TABLE 7 AROUND HERE]

Table 7 also reveals that migrants to high-income countries are on average much more educated than migrants to developing countries and the largest share of women is observed in North-North corridors. These patterns are better illustrated in Figure 3 which shows the gender and skill compositions of the 15 main corridors involving the South and North. Each bilateral corridor is represented by a grey bubble whose size is proportional to the migrant stock observed in 2000. Blank bubbles represent the average proportions of Table 7, and their size is proportional to the mean corridor size times 1,000, as each group includes many zeros and small corridors, the mean size is much lower than the size of the top-15 corridors in grey.

We have 7,614 South-North corridors with an average size of 5,959 migrants. The average proportion of college graduates is equal to 30.5 percent, and the share of women equals 46.8 percent. Looking across the main corridors, the skill shares vary from 6.3 percent for Turkey-Germany to 79.4 percent for India-US corridors. The share of women is usually between 40 and 60 percent with the exception of migration from Egypt and India to Saudi Arabia for which the share of women is around 14 percent.

We have 19,881 South-South corridors with an average size of 1,478 migrants. The average proportion of college graduates is equal to 10.4 percent, and the share of women is 48.9 percent. Looking at the main corridors, the skill ratio varies from 0.4 percent for Mali-Ivory Coast to 25.1 percent for Russia-Ukraine corridors. The share of women has a larger range of between 35 and 70 percent.

We have 2,916 North-North corridors with an average size of 7,702 migrants. The average proportion of college graduates is equal to 38.2 percent, and the share of women is 52.6 percent. The skill share varies from 4.0 percent for Portugal-France to 65.7 percent for UK-US corridors and the share of women is between 40 and 60 percent.

Finally, we have 7,614 North-South corridors with an average size of 434 migrants, a very small number relative to other cases. The average ratio of college graduates is equal to 21.5 percent, and the share of women is 47.7 percent. In the main corridors, the skill share ranges from 1.2 percent for Portugal-Mozambique to 38.8 percent for UK-South Africa. The share of women is again between 40 and 60 percent.

[INSERT FIGURE 3 AROUND HERE]

4 Conclusion

The literature on the causes and consequences of international migration has long been restrained by the paucity of harmonized, detailed and reliable data. In this paper, we take an important step towards resolving the situation and construct and characterize a comprehensive database documenting bilateral movements of high and low skilled workers by gender for 1990 and 2000. We use recorded bilateral migration data for a group of countries to estimate the determinants of migration which are then used to predict the missing data. We evaluate the precision of our predic-

tions and combine the recorded data with the most reliable estimates to construct comprehensive 195x195 migration matrices.

This database is clearly an evolving product that can be progressively improved by replacing estimates with official data or adding new census rounds. In its current state, our database provides reliable information on the origin, destination, gender and education level of about 100 million adult migrants in 2000, and 80 million in 1990. On this basis, we quantify migration stocks by region and income group, identify the main sources and destinations of human capital and the most important migration corridors. This database will prove to be key in understanding the demographic, economic. sociological and political implications of international migration.

5 References

Beine, M., F. Docquier and C. Ozden (2010): Diasporas. Journal of Development Economics, forthcoming.

Beine, M., F. Docquier and M. Schiff (2008): International migration. transfers of norms and home-country fertility. IRES Discussion Paper 2008-43.

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

Combes, P.-Ph., M. Lafourcade and Th. Mayer (2005): The trade creating effects of business and social networks: evidence from France. Journal of International Economics 66 (1): 1-29.

Docquier, F., E. Lodigiani, H. Rapoport and M. Schiff (2010): Emigration and the quality of home-country institutions, IRES Discussion Paper 2010-035.

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

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

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

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

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

Fargues, Ph. (2007): The demographic benefit of international migration: a hypothesis and its application to Middle Eastern and North African countries. In: Ozden. C. and M. Schiff (eds), International migration. economic development and policy. World Bank and Palgrave Macmillan: Washington DC.

Freedman, R., and D.P. Slesinger (1961): Fertility Differentials for the Indigenous Non-Farm Population of the United States. Population Studies 15(2): 161-173.

Goldberg, D. (1960): Another Look at the Indianapolis Fertility Data. Milbank Memorial Fund Quarterly 38 (1):23-36.

Gould, D. (1994): Immigrants Links to the Home Countries: Empirical Implication for U.S. bilateral Trade Flows. Review of Economics and Statistics 76 (2): 302-16.

Grogger, J., and G.H. Hanson (2008): Income maximisation and the selection and sorting of international migrants. NBER Working Paper n. 13821

Head, K., and J. Reis (1998): Immigration and Trade Creation: Econometric Evidence from Canada. Canadian Journal of Economics 31: 47-62.

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

Javorcik, B.S., C. Ozden, M. Spatareanu and I.C. Neagu (2010): Migrant Networks and foreign Direct Investment, Journal of Development Economics, forthcoming.

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

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

Minnesota Population Center (2010): Integrated Public Use Microdata Series. International: Version 6.0 [Machine-readable database]. Minneapolis: University of Minnesota.

Mohapatra, S., D. Ratha, and A. Silwal, (2010) Migration and Development Brief 13, World Bank

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

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

Pedersen, P.J., M. Pytlikova and N. Smith (2008): Selection or Network Effects? Migration Flows into 27 OECD Countries. 1990-2000. European Economic Review 52 (7): 1160-1186.

Peri, G., and F. Requena (2009): The Trade Creation Effect of Immigrants: Testing the Theory on the Remarkable Case of Spain. CReAM Discussion Paper Series n. 0915.

Pritchett, L. (2006): Let Their People Come: Breaking the Gridlock on Global Labor Mobility. Center for Global Development; University Presses Marketing.

Rauch, J.E., and A. Casella (2003): Overcoming informational barriers to international resource allocation: Prices and Ties. Economic Journal 113 (484): 21-42.

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

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

Spilimbergo, A. (2009). Democracy and foreign education. American Economic Review 99 (1): 528-43.

Winters., L., A., (2001): Assessing the Efficiency Gain from Further Liberalization: A Comment, in Suave, P., and Subramanian, A., (eds) Efficiency, Equity and Legitimacy: The Multilateral Trading System and the Millenium. Chicago: Chicago University Press. 106-113.

Walmsley, T. L. and L., A.,Winters,(2005): Relaxing the Restrictions on the Temporary Movement of Natural Persons: A Simulation Analysis, Journal of Economic Integration, December, 20(4).

6 Appendix

6.1 Migration data

Table A1 describes the various data sources and provides the total number of migrants together with the number of high-skill immigrants recorded in 1990 and 2000, by destination country.

[INSERT TABLE A1 ABOUT HERE]

6.2 Explanatory variables

Table A2 describes the data sources for the explanatory variables used in regressions of Section 3.2.

[INSERT TABLE A2 ABOUT HERE]

TABLES AND FIGURES

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

	OECD[a]	Non-OECD[b]	Imputed	*Imputed (%)*	Total
2000					
Males	28.170	11.856	11.478	*0.22*	51.504
High-skilled	10.310	1.903	1.583	*0.11*	13.796
Low-skilled	17.859	9.953	9.896	*0.26*	37.708
Females	29.209	8.354	11.464	*0.23*	49.027
High-skilled	10.021	1.156	1.215	*0.10*	12.393
Low-skilled	19.188	7.198	10.249	*0.28*	36.634
Total[c]	57.379	20.210	22.942	*0.23*	100.531
High-skilled	20.332	3.060	2.798	*0.11*	26.189
Low-skilled	37.047	17.151	20.144	*0.27*	74.342
1990					
Males	20.160	8.600	13.477	*0.32*	42.238
High-skilled	6.551	1.053	1.645	*0.18*	9.249
Low-skilled	13.610	7.547	11.832	*0.36*	32.989
Females	20.650	5.409	11.947	*0.31*	38.006
High-skilled	5.731	0.526	0.852	*0.12*	7.109
Low-skilled	14.920	4.883	11.095	*0.36*	30.897
Total[c]	40.811	14.009	25.424	*0.32*	80.244
High-skilled	12.281	1.580	2.498	*0.15*	16.358
Low-skilled	28.530	12.430	22.927	*0.36*	63.886

Notes. [a] 30 OECD destination countries in 1990 and 2000; [b] 31 non-OECD destinations in 1990 and 46 in 2000; [c] In order to obtain the exact numbers reported in DLM for OECD destinations, we need to add around 1.185 million migrants in 1990 and 1.240 in 2000 for which data by education levels are reported but country of origin data are missing (the corresponding numbers for non-OECD destinations are 1.292 and 0.715 million).

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

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

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

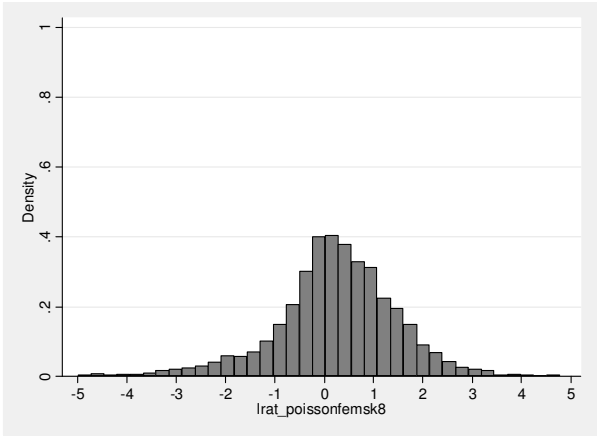
Table 3. In-sample comparison between actual and predicted stocks in 2000

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

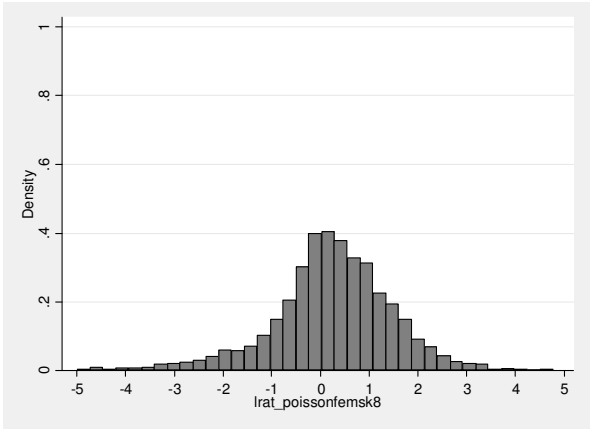
Notes: Column 'Obs' provides the number of observations; 'Corr' gives the correlation between our predictions and the actual migrant stocks; 'Migs' details the actual number of migrants (in millions); while 'As %' provides the percentage of the within sample migrants contained within each sample; 'Mean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

Figure 1. Distribution of the log-difference between actual and predicted stocks
High-skilled female migrants by corridor size in 2000

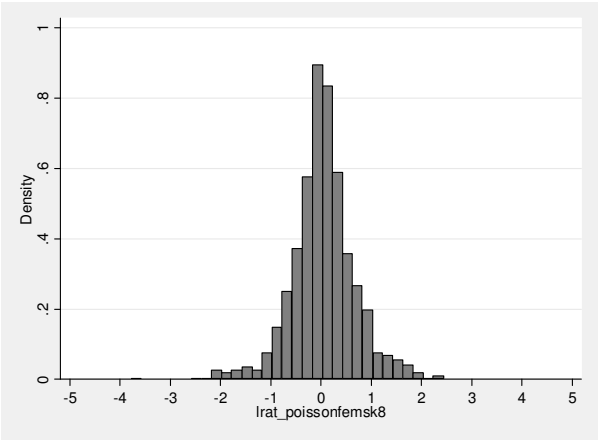
1a. All corridors



1b. Fewer than 250



1c. Greater than 250



1d. Greater than 5,000

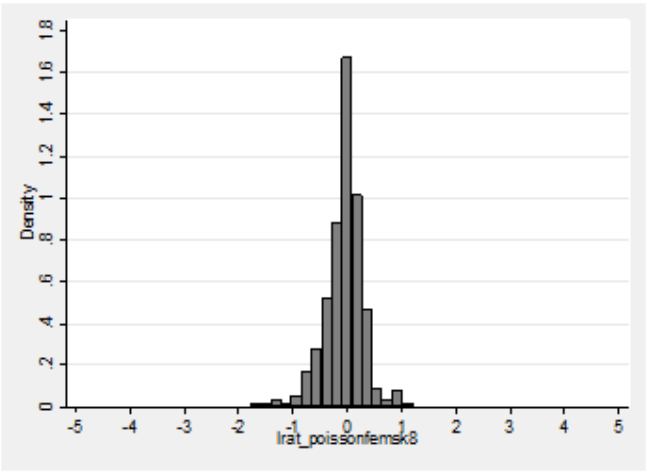


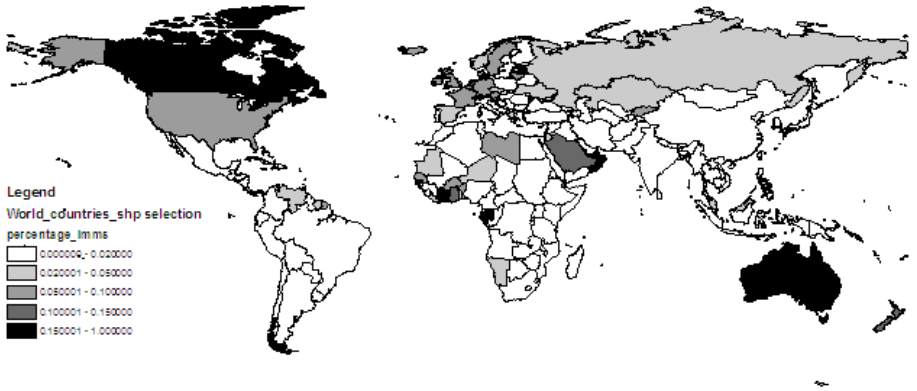
Table 4. Size (in millions), structure (in percent), and growth rate (in percent) of immigrant and emigrant stocks by region in 2000

	Immigration				**Emigration**			
Group	Total (mil.)	High-skill (%)	Women (%)	Growth (%)	Total (mil.)	High-skill (%)	Women (%)	Growth (%)
HIGH	67.8	33.0	48.7	36.1	25.8	36.0	52.0	9.1
DEV	32.7	11.6	48.8	7.6	74.8	22.6	47.7	32.0
LOW	*7.6*	*4.1*	*47.3*	*31.0*	*14.9*	*9.4*	*44.8*	*21.9*
LDC	*5.7*	*4.1*	*49.2*	*8.9*	*15.2*	*8.6*	*43.5*	*18.4*
SIDS	*0.9*	*24.2*	*52.2*	*38.4*	*4.5*	*34.3*	*55.1*	*57.2*
USA	24.3	42.7	51.0	62.1	0.9	58.8	50.4	18.9
CANZ	8.6	52.3	51.7	20.1	1.5	57.1	54.0	16.8
GCC	5.7	18.8	20.9	24.4	0.4	22.9	35.4	23.4
EU27	22.3	21.9	50.5	26.2	20.2	31.5	51.9	6.4
LAC	2.4	15.7	49.1	32.1	15.6	25.0	50.2	89.5
SSA	12.7	2.7	45.5	21.7	14.1	8.3	45.6	28.8
CIS	7.9	19.0	55.0	5.7	10.6	24.1	55.8	5.7
INDIA	1.6	10.0	48.8	-26.3	4.9	31.8	39.7	2.8
CHINA	0.1	36.4	56.0	237.3	4.1	28.1	52.4	25.9
MENA	4.2	22.5	49.2	4.0	8.4	23.8	38.5	24.2
ASIA	8.2	18.3	51.1	6.0	16.2	32.2	48.6	33.9
OTHERS	2.4	20.9	52.8	19.1	3.6	22.3	49.0	30.8
TOTAL	**100.5**	**26.1**	**48.8**	**25.3**	**100.5**	**26.1**	**48.8**	**25.3**

Notes. Colomn 'Total' gives the aggregate regional stock of migrants in millions (mil.) ; 'High-skill' gives the percentage (if %) or stock (if mil.) of high-skilled migrants ; 'Women' gives the percentage (if %) or stock (if mil.) of female migrants ; and 'Growth' gives the growth rate of the migration stock over 1990-2000. 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, ASIA : Other Asian countries excluding GCC, CIS and MENA. Each country only belongs to one geographical group.

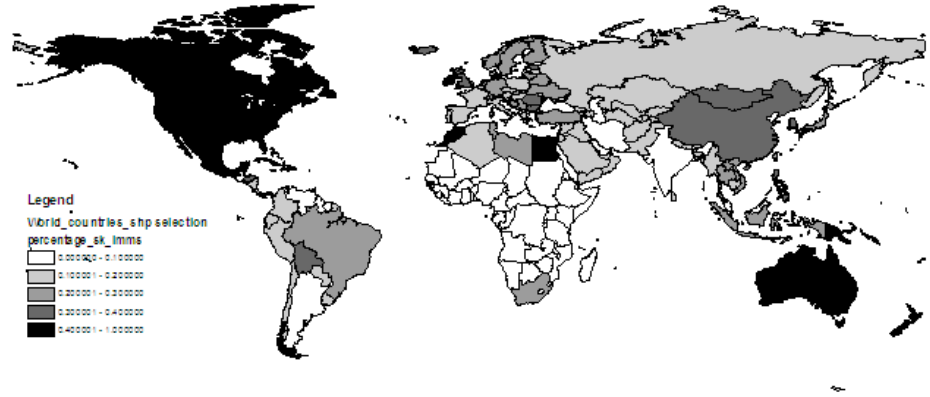
Figure 2. Geographical distribution of migration intensity and positive selection in 2000

2a. Immigration stock as percentage of the population



2b. Emigration stock as percentage of the population

2c. High-skilled immigration as percentage of total immigration



2d. High-skilled emigration as percentage of total emigration

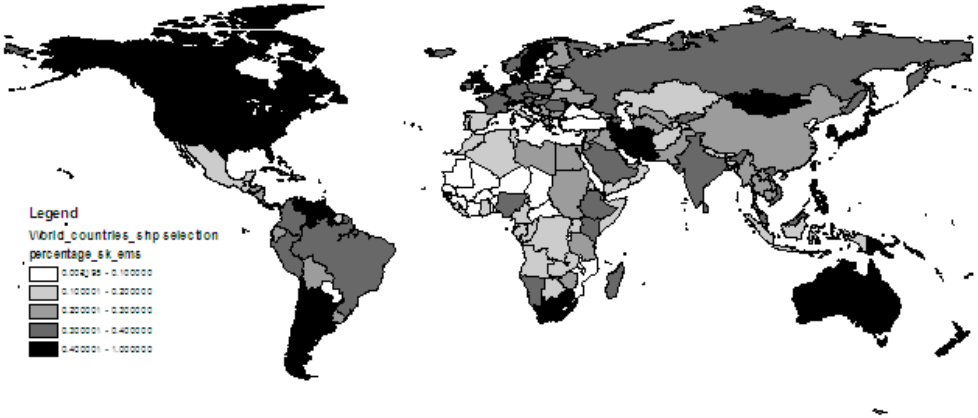


Table 5. Main receiving and sending countries in 2000

Total migration stock (in millions)					**High-skilled migration stock (in millions)**				**Migration stock as percentage of population (%)**				
Main receiving countries					Main receiving countries				Main receiving countries				
Country	Immig.	*High-sk*[a]	*Women*[a]	*Growth*[a]	Country	Immig.	*Women*[b]	*Growth*[b]	Country	Immig.	*High-sk*[a]	*Women*[a]	*Growth*[a]
U.S.	24.3	*42.7*	*51.0*	*62.1*	U.S.	10.3	*50.0*	*66.8*	Macao	81.3	*16.2*	*60.7*	*13.1*
Russia	4.9	*18.1*	*54.3*	*-2.2*	Canada	2.7	*48.4*	*45.2*	Gabon	67.8	*1.5*	*44.5*	*-8.8*
Germany	4.7	*21.8*	*46.2*	*45.9*	Australia	1.6	*49.3*	*39.5*	Qatar	67.0	*18.3*	*25.8*	*27.3*
Canada	4.6	*58.8*	*52.3*	*25.2*	U.K.	1.2	*50.2*	*122.8*	Cote d'Ivoire	65.8	*0.9*	*44.7*	*19.6*
Cote d'Ivoire	3.9	*0.9*	*44.7*	*19.6*	Germany	1.0	*45.2*	*87.4*	West B-Gaza	63.9	*9.8*	*48.9*	*25.4*
France	3.7	*16.4*	*49.5*	*8.8*	Russia	0.9	*52.2*	*58.8*	Gambia	60.3	*0.5*	*48.1*	*97.4*
U.K.	3.5	*34.9*	*53.0*	*29.7*	France	0.6	*46.6*	*80.6*	U.A.E.	59.9	*18.4*	*28.0*	*71.8*
Australia	3.5	*45.2*	*50.8*	*12.9*	Saudi Arabia	0.6	*11.2*	*45.2*	Kuwait	50.8	*19.2*	*32.8*	*36.6*
Saudi Arabia	3.1	*18.8*	*14.1*	*8.3*	Israel	0.5	*53.2*	*220.1*	Israel	45.4	*33.9*	*54.5*	*28.1*
Ghana	2.1	*1.6*	*39.3*	*80.8*	Netherlands	0.4	*51.6*	*46.6*	Bahrain	41.5	*20.8*	*44.0*	*32.7*
Main sending countries					Main sending countries				Main sending countries				
Country	Emig.	*High-sk*[a]	*Women*[a]	*Growth*[a]	Country	Emig.	*Women*[b]	*Growth*[b]	Country	Emig.	*High-sk*[a]	*Women*[a]	*Growth*[a]
Mexico	6.5	*14.8*	*45.4*	*138.6*	U.K.	1.6	*47.1*	*31.5*	Suriname	89.4	*18.4*	*53.5*	*19.3*
India	4.9	*31.8*	*39.7*	*2.8*	India	1.6	*36.9*	*69.3*	Burkina F.	82.7	*0.3*	*45.8*	*23.5*
China	4.1	*28.1*	*52.4*	*25.9*	Philippines	1.2	*59.7*	*80.1*	Guyana	69.2	*40.2*	*54.9*	*72.3*
U.K.	3.4	*46.6*	*51.0*	*4.2*	China	1.1	*48.5*	*98.7*	Cape Verde	57.4	*8.3*	*53.6*	*55.0*
Ukraine	3.3	*22.2*	*58.2*	*-14.1*	Germany	1.0	*51.6*	*19.6*	Jamaica	53.9	*42.1*	*57.0*	*47.5*
Burkina F.	2.9	*0.3*	*45.8*	*23.5*	Mexico	1.0	*47.2*	*159.7*	Togo	52.2	*1.8*	*40.6*	*45.9*
Russia	2.8	*30.9*	*57.5*	*-8.9*	Russia	0.9	*55.9*	*47.2*	Barbados	48.9	*41.3*	*54.5*	*18.1*
Italy	2.6	*16.5*	*46.7*	*-4.0*	Ukraine	0.7	*55.2*	*90.5*	Mauritania	43.2	*1.2*	*22.7*	*126.7*
Germany	2.6	*39.6*	*56.8*	*4.3*	Korea	0.6	*52.1*	*84.4*	West B-Gaza	41.8	*20.1*	*45.7*	*4.7*
Philippines	2.3	*54.5*	*59.0*	*66.9*	Vietnam	0.5	*45.5*	*135.4*	Guinea-Bissau	39.6	*1.1*	*63.6*	*214.4*

Notes. We exclude microstates with less than 100,000 inhabitants. Superscript [a] denotes percentage or 1990-2000 growth rate of the total migration stock. Superscript [b] denotes percentage or 1990-2000 growth rate of the high-skill migration stock.

Table 6. Main migration corridors in 2000 and in 1990-2000 (top-40)

Largest migration stocks in 2000 (in thousands)					Largest migration net flows 1990-2000 (in thousands)				
Origin	Destination	Stock	*H-S*	*Women*	Origin	Destination	Flow	*H-S*	*Women*
Mexico	U.S.	6,374.8	*14.4*	*45.2*	Mexico	U.S.	3,718.8	*15.2*	*45.1*
Burkina F.	Ivory C.	2,238.5	*0.3*	*44.7*	India	U.S.	490.6	*81.2*	*46.0*
Ukraine	Russia	1,926.1*	*15.8*	*58.3*	Vietnam	U.S.	484.9	*44.2*	*51.9*
China	Hong Kong	1,620.5	*12.9*	*50.6*	Kazakhstan	Russia	467.6*	*22.4*	*52.5*
Turkey	Germany	1,272.0	*6.3*	*45.8*	Philippines	U.S.	418.2	*80.8*	*62.5*
Philippines	U.S.	1,163.6	*71.7*	*58.6*	Turkey	Germany	390.0	*9.2*	*47.6*
Kazakhstan	Russia	1,069.6*	*18.8*	*51.1*	China	U.S.	378.7	*57.9*	*54.3*
Russia	Ukraine	988.9*	*25.1*	*56.0*	Burkina F.	Ivory C.	367.5	*0.3*	*44.7*
U.K.	Australia	969.0	*39.4*	*50.0*	El Salvador	U.S.	341.3	*19.4*	*48.3*
India	Pakistan	905.6*	*14.1*	*46.5*	Ivory C.	Burkina F.	323.8*	*0.7*	*56.3*
Bangladesh	India	902.5*	*7.8*	*47.8*	Dom. Rep.	U.S.	287.0	*23.0*	*56.0*
China	U.S.	841.7	*51.6*	*52.6*	Korea	U.S.	265.5	*78.1*	*57.1*
India	U.S.	836.8	*79.4*	*45.2*	Russia	Kazakhstan	250.0*	*21.4*	*60.1*
Vietnam	U.S.	807.3	*43.0*	*50.8*	Togo	Ghana	242.8*	*0.9*	*1.4*
Cuba	U.S.	803.5	*38.3*	*51.7*	Mauritania	Ghana	211.6*	*0.8*	*13.9*
Mali	Ivory C.	792.3	*0.4*	*45.0*	Guatemala	U.S.	204.3	*17.3*	*44.6*
Belarus	Russia	717.1*	*14.7*	*57.4*	Jamaica	U.S.	203.7	*46.4*	*60.1*
Canada	U.S.	715.8	*61.4*	*56.7*	China	Hong Kong	198.6*	*51.7*	*59.7*
Korea	U.S.	676.6	*64.6*	*58.3*	Colombia	U.S.	181.7	*44.9*	*58.5*
Togo	Ghana	651.5*	*1.0*	*36.1*	Burkina F.	Ghana	180.9*	*0.1*	*32.1*
Germany	U.S.	646.8	*59.8*	*60.4*	Haiti	U.S.	168.1	*46.6*	*59.2*
U.K.	U.S.	637.6	*65.7*	*56.3*	India	U.A.E	156.2	*23.9*	*23.3*
El Salvador	U.S.	619.2	*18.3*	*50.4*	China	Canada	142.3	*65.9*	*54.6*
India	Saudi Arabia	604.6	*18.8*	*14.1*	Uzbekistan	Russia	137.6*	*21.2*	*57.1*
Egypt	Saudi Arabia	588.1	*18.8*	*14.1*	Cuba	U.S.	135.9	*55.3*	*49.1*
U.K.	Canada	580.3	*63.0*	*53.5*	Russia	Kyrgyzstan	133.9*	*12.9*	*64.5*
Portugal	France	536.2	*4.0*	*48.6*	Honduras	U.S.	133.7	*21.0*	*48.1*
Dom. Rep.	U.S.	527.5	*27.3*	*55.8*	Mali	Ivory C.	130.1	*0.4*	*45.0*
Burkina F.	Ghana	517.2*	*0.3*	*47.8*	Ukraine	Israel	128.9	*66.7*	*56.7*
Algeria	France	512.8	*10.2*	*43.1*	Ecuador	U.S.	128.1	*28.9*	*47.3*
Italy	U.S.	461.1	*28.7*	*51.3*	India	Canada	124.1	*59.4*	*51.3*
Italy	Germany	456.0	*9.6*	*40.5*	Pakistan	U.A.E	124.1	*23.5*	*18.1*
Jamaica	U.S.	449.8	*44.3*	*57.2*	Philippines	Hong Kong	120.3*	*25.7*	*95.7*
Indonesia	Malaysia	437.0*	*1.3*	*42.7*	Morocco	Spain	119.6	*11.1*	*38.2*
Morocco	France	423.8	*12.8*	*45.0*	Tajikistan	Russia	115.3*	*20.7*	*48.6*
Ireland	U.K.	422.4	*24.6*	*54.8*	Paraguay	Argentina	114.6	*2.4*	*61.4*
Pakistan	India	418.1*	*12.6*	*50.4*	Peru	U.S.	112.8	*50.2*	*51.6*
Ivory C.	Burkina F.	406.7*	*1.4*	*54.3*	Vietnam	Cambodia	112.6*	*11.1*	*56.2*
Colombia	U.S.	402.9	*45.8*	*56.7*	Italy	Germany	106.0	*24.5*	*42.9*
Korea	Japan	400.7	*23.0*	*53.8*	China	Japan	104.2	*36.7*	*62.4*

Notes. Migration net flows 1990-2000 = Migration stock in 2000 minus Migration stock in 1990. Columns 'H-S' and 'Women' give the proportion of high-skilled migrants and women in total migration. A superscript * denotes imputed value.

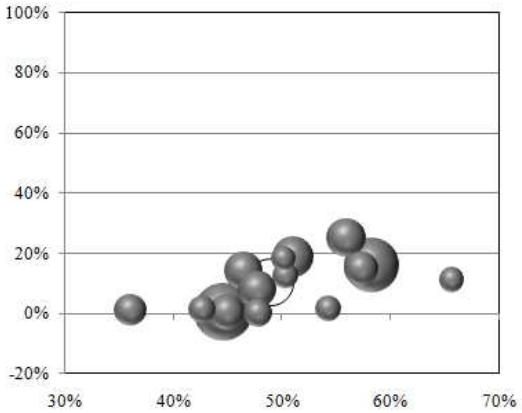
Table 7. Migration between North and South countries in 2000

		High-income (destination)	Developing (destination)
High-income (origin)	Stock (in millions)	22.5	3.3
	High-skill (percent)	*38.2*	*21.5*
	Women (percent)	*52.6*	*47.7*
	Growth (percent)	*8.6*	*12.5*
Developing (origin)	Stock (in millions)	45.4	29.4
	High-skill (percent)	*30.5*	*10.4*
	Women (percent)	*46.8*	*48.9*
	Growth (percent)	*55.6*	*7.0*

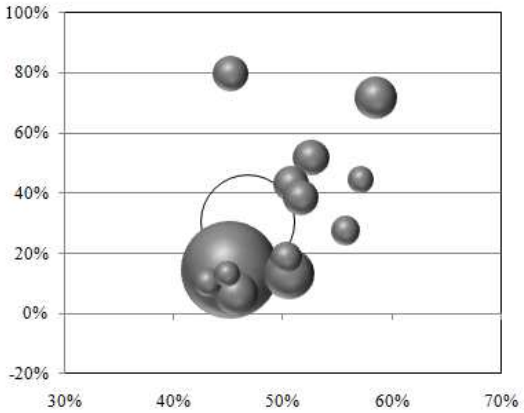
Notes. For the definition of high-income and developing countries. we use the World Bank classification.

Figure 3. Gender (X-axis) and skill (Y-axis) compositions of the main migration corridors per income group (Year 2000)

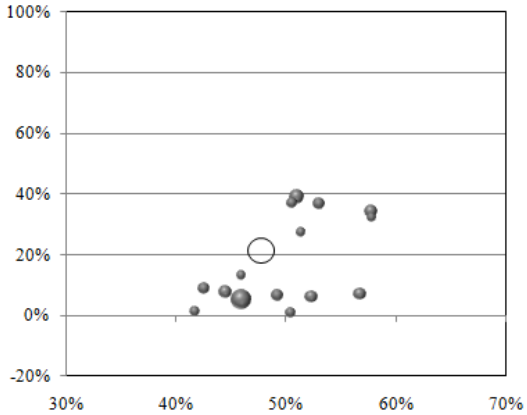
3a. South-South corridors



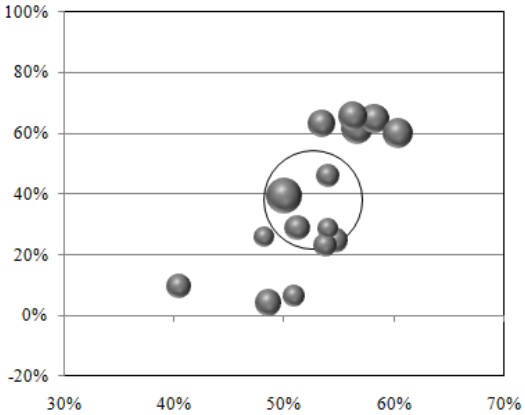
3b. South-North corridors



3c. North-South corridors



3d. North-North corridors



Note. Fifteen main corridors by income group are represented by grey bubbles, with bubble's surface proportional to the corridor size. Blank bubbles represent mean corridors per income group, with bubble's surface proportional to the mean size times 1,000. Share of women is represented on the horizontal axis ; Share of high-skilled is represented on the vertical axis ; 'North' = High-income countries ; 'South' = Developing countries.

Table A1. Migration data to selected non-OECD destinations

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

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

Table A2. Description of Explanatory Variables

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

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