

Income Maximization and the Selection and Sorting of International Migrants

Jeffrey Grogger, University of Chicago and NBER

Gordon H. Hanson, UCSD and NBER

July 2008

Abstract. Two prominent features of international labor movements are that the more educated are more likely to emigrate (positive selection) and more-educated migrants are more likely to settle in destination countries with high rewards to skill (positive sorting). Using data on emigrant stocks by schooling level and source country in OECD destinations, we find that a simple model of income maximization can account for both phenomena. Results on selection show that migrants for a source-destination pair are more educated relative to non-migrants the larger is the absolute skill-related difference in earnings between the destination country and the source. Results on sorting indicate that the relative stock of more-educated migrants in a destination is increasing in the absolute earnings difference between high and low-skilled workers. We use our framework to compare alternative specifications of international migration, estimate the magnitude of migration costs by source-destination pair, and assess the contribution of wage differences to how migrants sort themselves across destination countries.

We thank for comments Eli Berman, George Borjas, Gordon Dahl, Frederic Docquier, Hillel Rapoport, Dean Yang, and seminar participants at the University of Chicago, Federal Reserve Bank of Atlanta, Harvard, LSE, Princeton, UCSD, UC Irvine, UCL, Yale, and the 2007 NBER Summer Institute. Any errors are ours alone.

1. Introduction

International migration is on the rise. As of 2005, 3% of the world's population resided outside their country of birth. Most migrants leave home bound for rich nations. The UN (2006) estimates that in 2005, 40.9% of the global emigrant population resided in just eight rich economies, with 20.2% living in the US alone. In major destination countries, the number of foreign born is growing, reaching 12.5% of the total population in the US, 11.2% in Germany, 10.5% in France, and 8.2% in the UK.¹

One striking feature of international labor flows is that the more educated are those most likely to move abroad. Using data from Docquier and Marfouk (2006), Figure 1 plots the share of tertiary-educated emigrants against the share of tertiary-educated non-emigrants by source country. Emigrants are generally positively selected in terms of schooling, in that they are more educated than their non-migrant counterparts. This observation has renewed interest in brain drain from developing economies.²

A second – and perhaps less appreciated – feature of international migration is the sorting of emigrants across destinations. Countries with large absolute skill-related wage differences attract a disproportionate share of more-educated emigrants. Table 1, also based on data from Docquier and Marfouk (2006), gives the share of international migrants residing in OECD countries by major destination region. The US and Canada, where skill-related earnings differences are relatively large, receive 51.4% of the OECD's immigrants, but 65.5% of its immigrants with tertiary schooling. Europe, where

¹ In 1990, these figures were 7.9% in the US, 7.3% in Germany, 6.3% in France, and 3.2% in the UK, indicating that a substantial share of immigrants in these countries have arrived recently.

² See Adams (2003), Beine, Docquier and Rapoport (2001, 2007, 2008), Docquier and Rapoport (2007), Kapur and McHale (2005), and Docquier, Lohest, and Marfouk (2007).

skill-related earnings differences are relatively small, receives 38.4% of the OECD's immigrants, but only 23.6% of its tertiary-schooled immigrants.

In this paper, we estimate a simple model of international migration using the Roy (1951) income maximization framework. The Roy model is the basis for a large body of work on migration (Borjas, 1999). In our application, an increase in the absolute difference in earnings between high and low skilled workers in destination countries causes migration from source countries to rise and the mix of migrants to become more skilled. Using data from Beine, Docquier and Rapoport (2007) on the stock of migrants by education level and by source country residing in OECD destinations as of 2000, we find strong support for the income maximization hypothesis.

A growing literature on global labor flows considers the causes of emigration from developing countries (Docquier, Lohest, and Marfouk, 2007), how average income and income inequality affect bilateral migration flows (Mayda, 2005; Clark, Hatton, and Williamson, 2007), and the impact of labor productivity and log earnings inequality on skilled migration (Rosenzweig, 2007; Brücker and Defoort, 2006; Belot and Hatton, 2008).³ Much of this work is also based on the Roy model. Most of these studies focus on migrant selectivity, although some also address the scale of migration.

Our paper provides an integrated framework to examine the scale of migration, the selectivity of migrants in terms of schooling, and the sorting of migrants across destinations. We use the framework to identify sources of misspecification in estimating the Roy model, address the nature of migrant selectivity, and evaluate the contribution of wage differences to the concentration of skilled migrants in particular destinations.

Our model delivers separate estimating equations for scale, selectivity, and sorting

³ See also Volger and Rotte (2000), Pedersen, Pytlikova and Smith (2004), Hatton and Williamson (2005).

in migration. The model predicts that earnings should have the same effect on each, but the specifications differ in the data they require. The scale regression calls for earnings by schooling level in the source and destination and for data on fixed migration costs. The selection regression, in contrast, differences out fixed migration costs. While results for both specifications support income maximization, they reveal that by not controlling for unobserved migration costs in the scale regression, we severely underestimate the impact of earnings on migration. This finding may account for the unstable relationship between income and migration in the literature (e.g., Mayda, 2005). Our multilateral results on wage differences and selection are consistent with Rosenzweig (2007), who examines legal migration to the US and finds that source-country emigration rates are decreasing in source-country labor productivity.

The sorting regression controls for all source-specific determinants of migration, which means we can estimate the impact of earnings on migration without source-country wage data. This is an advantage, as our source-country wage measures are coarser than the survey-based micro-data that is available to estimate destination-country wages. Reassuringly, sorting regressions are similar for the two types of wage measures. The sorting and selection regressions are similar as well, as required by the theory.

The sorting regressions also allow us to examine the effects of taxes. Post-tax earnings are a stronger correlate of migration than pre-tax earnings, consistent with migrants weighing tax treatment. Previous studies of sorting focus on internal US migration (Borjas, Bronars, and Trejo, 1992; Dahl, 2002), and are silent on tax effects. Work on international migration tends to focus on selection and ignore sorting.

A further contribution of our paper is to address conflicting results on migrant

selectivity. In seminal work, Borjas (1987) develops a version of the Roy model which predicts that migrants who move from a source country with high returns to skill to a destination with low returns to skill should be negatively selected. Although the Borjas framework performs well in explaining migration from Puerto Rico to the US (Ramos, 1992; Borjas, 2006), it does less well in explaining Mexico-to-US migration, where migrants are drawn from the middle of the skill distribution.⁴ Other work has tested the Roy model by estimating the correlation between the skill composition of bilateral migration and log income or log wage inequality. Those results are sensitive to the countries included in the sample (Brücker and Defoort, 2006; Belot and Hatton, 2008).

To reconcile the Roy model with the strong positive selectivity seen in Figure 1, we compare two models. Our base specification explains migration with *absolute* wage differences, which is consistent with linear utility. An alternative specification explains migration with *relative* wage differences, which is consistent with Borjas (1987) and with log utility. The data strongly reject log utility, implying that migration responds to absolute, not relative, rewards to skill. Since absolute skill-related earnings differences are much larger in destination than in source countries, positive selection of migrants is a logical outcome. We show further that migrant selection and migrant sorting are independent of each other. While the selectivity of migrants by skill depends on the reward to skill in the source country, the sorting of migrant by skill does not. Positive sorting is a general implication of income maximization.

In applications of our results, we use the selection regression to estimate bilateral migration costs between source and destination countries, finding that these are often an

⁴ See Chiquiar and Hanson (2005), Orrenius and Zavodny (2005), McKenzie and Rapoport (2006), Ibarrran and Lubotsky (2005), and Fernandez-Huertas (2006).

order of magnitude larger than source-country wages for low-skilled workers. We then use the sorting regression to decompose differences in immigrant skills across destination countries into components due to wage differences, language, distance, and other factors. Skill-related wage differences are the dominant factor in explaining why the US and Canada receive more skilled immigrants than other OECD destinations.

2. Theory and Empirical Specification

A. A Model of Scale, Selection and Sorting in Migration

Consider migration flows between many source countries and many destination countries. To be consistent with our data, assume that workers fall into one of three skill groups, corresponding to primary, secondary, or tertiary education. Let the wage for worker i with skill level j from source country s in destination country h be⁵

$$(1) \quad W_{ish}^j = \exp(\mu_h + \delta_h^2 D_{is}^2 + \delta_h^3 D_{is}^3),$$

where $\exp(\mu_h)$ is the wage for workers with primary education, δ_h^2 (δ_h^3) is the return to secondary (tertiary) education, and $D_{is}^j = 1$ if person i from source s has schooling level j .

Let C_{ish}^j be the cost of migrating from s to h for worker i with skill level j , which we assume to have two components: a fixed monetary cost of moving from s to h , f_{sh} ; and a component that varies by skill, g_{sh}^j (which may be positive or negative), such that

$$(2) \quad C_{ish}^j = f_{sh} + g_{sh}^1 D_i^1 + g_{sh}^2 D_i^2 + g_{sh}^3 D_i^3.$$

Migration costs are influenced by the linguistic and geographic distance between the source and the destination and by destination-country immigration policies. The impacts

⁵ In (1), we do not allow for unobserved components of skill that may affect wages, which are of central concern in Borjas (1987, 1991). Since our data on migrant stocks are aggregated by level of schooling and source country, it is not possible to address within education group heterogeneity in skill.

of these characteristics may depend on the migrant's skill due to time costs associated with migration or skill-specific immigration policies in the destination.

Our primary interest is in a linear utility model where the utility associated with migrating from country s to country h is a linear function of the difference between wages and migration costs as well as an unobserved idiosyncratic term ε_{ish}^j such that

$$(3) \quad U_{ish}^j = \alpha(W_{ih}^j - C_{ish}^j) + \varepsilon_{ish}^j,$$

where $\alpha > 0$. We think of (3) as a first-order approximation to some general utility function, with the marginal utility of income given by α . One of the “destinations” is the source country itself, for which migration costs are zero.

Assuming that workers choose whether and where to emigrate so as to maximize their utility, and assuming that ε_{ish}^j follows an i.i.d. extreme value distribution, we can apply the results in McFadden (1974) to write the log odds of migrating to destination country h versus staying in the source country s for members of skill group j as⁶

$$(4) \quad \ln \frac{E_{sh}^j}{E_s^j} = \alpha(W_h^j - W_s^j) - \alpha f_{sh} - \alpha g_{sh}^j$$

where E_{sh}^j is the population share of education group j in s that migrates to h , E_s^j is the population share of education group j in s that remains in s , and $W_h^j = e^{\mu_h + \delta_h^j}$. Equation

(4) speaks to the scale of migration. It says that income maximization, together with our

⁶ The specification of the disturbance in equation (3) embodies the assumption that IIA applies among destination countries. In the empirical analysis, the sample of destination countries is limited to OECD members. To use (4) as a basis for estimation, we need only that IIA applies to the OECD countries in the sample. The analysis is thus consistent with a more complicated nesting structure, in which we examine only the OECD branch of the decision tree. Alternatively, one might imagine that there are multiple branches of the decision tree even among OECD destinations, such that IIA fails. In the estimation, we test for this possibility, following the logic of Hausman and McFadden (1984).

assumptions about utility and the error terms, implies that the skill-group-specific log odds of migrating to h from s should depend positively on the level difference in skill-specific wages between h and s and negatively on migration costs.

To analyze emigrant selection, take the difference of equation (4) between tertiary- and primary-educated workers to yield:

$$(5) \quad \ln \frac{E_{sh}^3}{E_{sh}^1} - \ln \frac{E_s^3}{E_s^1} = \alpha[(W_h^3 - W_s^3 - g_{sh}^3) - (W_h^1 - W_s^1 - g_{sh}^1)].$$

The first term on the left of (5) is a measure of the skill mix of emigrants from source s to destination h , which we refer to as the log skill ratio. The numerator is the share of tertiary-schooled workers in s who migrate to h ; the denominator is the share of primary-schooled workers in s who migrate to h . The second term on the left of (5) is the log skill ratio for non-migrants in s , meaning the full expression on the left of (5) is the difference in skill mix between emigrants (from s to h) and non-migrants for source country s .

If the left side of (5) is negative, emigrants are negatively selected; if it is positive, they are positively selected. Since $\alpha > 0$, equation (5) indicates that emigrants should be positively selected if the wage difference between the source and destination, net of skill-varying migration costs, is greater for high-skill workers. Note that differencing between skill groups eliminates fixed costs f_{sh} from the selection equation in (5).

To analyze the model's implications for how emigrants should sort themselves across destinations, collect those terms in (5) that vary only by source country to yield

$$(6) \quad \ln \frac{E_{sh}^3}{E_{sh}^1} = \alpha(W_h^3 - W_h^1) - \alpha(g_{sh}^3 - g_{sh}^1) + \tau_s$$

where $\tau_s = \ln(E_s^3 / E_s^1) - \alpha(W_s^3 - W_s^1)$. Fixed costs do not appear in the sorting equation (6) because they are absent from the selection equation (5).

Equation (6) states that emigrants from a given source country should sort themselves across destinations by skill according to the rewards to skill in different destinations. If the (net) rewards to skill are higher in destination h than in destination k, then destination h should receive a higher-skilled mix of emigrants from source country s than should destination country k. Put differently, higher skill-related wage differences should give destination countries an advantage in competing for skilled immigrants.

B. Relationship to Earlier Research

The model summarized in (4), (5), and (6) highlights the role of fixed costs and *absolute* wage differences in influencing the scale, selectivity, and sorting of migration flows. In contrast, much of the literature focuses on *relative* returns to skill and assumes migration costs are proportional to income (see Borjas, 1991 and 1999). To compare these two frameworks, consider a model where wages and migration costs are as before, but utility is logarithmic and the error term is proportional rather than additive:

$$(7) \quad U_{ish}^j = (W_{ih}^j - C_{ish}^j)^\lambda \exp(v_{ish}^j)$$

We assume $\lambda > 0$ and v_{ish}^j follows an i.i.d. extreme value distribution. The analogues to the scale, selection and sorting equations in (4), (5), and (6) for this model are given by

$$(8) \quad \ln \frac{E_{sh}^j}{E_s^j} = \lambda(\ln W_h^j - \ln W_s^j) - \lambda m_{sh}^j$$

$$(9) \quad \ln \frac{E_{sh}^3}{E_{sh}^1} - \ln \frac{E_s^3}{E_s^1} = \lambda(\delta_h^3 - \delta_s^3) - \lambda(m_{sh}^3 - m_{sh}^1)$$

$$(10) \quad \ln \frac{E_{sh}^3}{E_{sh}^1} = \lambda \delta_h^3 - \lambda (m_{sh}^3 - m_{sh}^1) + \rho_s$$

where $m_{sh}^j = (f_{sh} - g_{sh}^j) / W_h^j$ and $\rho_s = \ln(E_s^3 / E_s^1) - \lambda \delta_s^3$.⁷ In the log utility model, the scale of migration is influenced by the relative wage difference between the source and destination countries (see (8)), and selectivity and sorting are functions of returns to skill, as given by the δ terms, rather than skill-related level wage differences (see (9) and (10)).

With log utility, differencing between skill groups does not in general eliminate migration costs from the selection or sorting equations in (9) and (10). Where skill-varying costs are proportional to wages, such that $g_{sh}^j = \pi_{sh} W_h^j$, differencing between skill groups eliminates skill-varying costs, but not fixed costs. Since much of the literature focuses on models where skill-varying costs are assumed to be proportional to wages and fixed costs are assumed to be zero, it is a case of special interest.

Examining conditions for migrant selectivity provides a useful way of comparing our linear utility model with fixed migration costs to the more standard log utility model with proportional migration costs. To analyze our linear utility model, substitute (1) into (5), rearrange terms, and make use of the fact that $e^\delta - 1 \approx \delta$. Our linear utility model then predicts that emigrants should be negatively selected in terms of skill if

$$(11) \quad \frac{\delta_s^3}{\delta_h^3} > \left[\frac{W_h^1}{W_s^1} \left(1 + \frac{g_{sh}^3}{(W_s^3 - W_s^1)} \right)^{-1} \right].$$

⁷ In deriving (8), we use the approximation that $\ln(W-C) \approx \ln W - C/W$ for sufficiently small C/W . Equation (9) follows from the fact that $\ln W_h^3 - \ln W_h^1 = \delta_h^3$. Belot and Hatton (2008) use the Roy model to motivate an empirical specification similar to equation (9).

In the case where $g_{sh}^3 = 0$ (i.e., fixed migration costs are independent of skill), the condition for negative selection is $\delta_s^3 / \delta_h^3 > W_h^1 / W_s^1$. Under log utility, zero fixed costs, and skill-varying costs that are proportional to wages, equation (9) shows that negative selection will obtain if $\delta_s^3 / \delta_h^3 > 1$, as in Borjas (1987).

With north-to-north migration similar productivity levels between the source and the destination imply that low-skill wages are also similar, such that $W_h^1 \approx W_s^1$. In that case, both models predict that emigrants who move from a source with high returns to skill to a destination with low returns should be negatively selected. However, the models make different predictions in the context of much south-to-north migration, where differences in productivity imply that $W_h^1 \gg W_s^1$. Here, our linear utility model predicts negative selection only when the relative return to skill in the source country (δ_s^3 / δ_h^3) exceeds the relative productivity advantage of the destination country (W_h^1 / W_s^1).⁸

To preview the data, Figure 2 plots the log odds of emigrating for the tertiary educated against that for the primary educated by source country, which is the dependent variable in (5) and (9). Nearly all points are above the 45 degree line, indicating positive selection. Figure 3a plots the log ratio of incomes at the 80th and 20th percentiles against log income at the 20th percentile for the sample countries.⁹ The series are negatively correlated, suggesting poor countries have larger *relative* differences in income between the high and low skilled. This is consistent with evidence that the returns to schooling

⁸ Factoring in skill-specific migration costs makes predictions about selection even more ambiguous in the linear utility/fixed cost model. Recall that skill specific costs in (11), g_{sh}^3 , may be positive or negative. If more skilled workers tend to have higher (lower) costs, the likelihood of negative selection would be higher (lower) than the base case of no skill-specific costs.

⁹ We discuss construction of income measures in the section 3.

tend to be higher in developing than in developed countries (Psacharopoulos et al., 2004; Hanushek and Zhang, 2006). Since most destination countries are clustered in the lower right of the graph, the log utility/proportional cost model predicts emigrants from developing countries should be negatively selected, clearly at odds with Figure 2.

Figure 3b plots the level difference in incomes between the 80th and 20th percentiles against the level of income at the 20th percentile. The strong positive relation indicates that rich countries tend to have larger *absolute* income differences between the high and low skilled. With most destination countries in the upper right of the graph, linear utility implies that emigrants should be positively selected, which concords with Figure 2. These graphs are initial evidence the data favor linear utility over log utility.

While many studies test for migrant selectivity, fewer examine migrant sorting across multiple destinations. Borjas, Bronars, and Trejo (1992) develop a theoretical model of sorting on the basis of destination returns to skill. They and Dahl (2002) estimate empirical models of sorting using data on internal US migration. One point that seems to have escaped the literature is that selection and sorting are independent. In terms of our model, sorting between destinations h and k depends on the sign of

$$\Delta^{hk} = [W_h^3 - W_h^1 - (g_{sh}^3 - g_{sh}^1)] - [W_k^3 - W_k^1 - (g_{sk}^3 - g_{sk}^1)],$$

whereas from (5), selection to destination h depends on the sign of

$$\Delta^h = (W_h^3 - W_s^3 - g_{sh}^3) - (W_h^1 - W_s^1 - g_{sh}^1).$$

Since selection depends on source-country wages, whereas sorting does not, sorting is independent of selection. If $\Delta^{hk} > 0$, destination h should receive more highly skilled migrants than destination k. This holds whether emigrants from s to h and k exhibit selectivity that is positive ($\Delta^h > 0, \Delta^k > 0$), negative ($\Delta^h < 0, \Delta^k < 0$), or bimodal.

C. Estimation

The models we develop make it straightforward to derive estimating equations which can be used to test for income maximization. Let x_{sh} be a vector of characteristics of the source-destination pair, such as geographic and linguistic distance, and let skill-varying costs be given by $g_{sh}^j = x_{sh}\theta^j$. The empirical version of the scale equation is

$$(12) \quad \ln \frac{\hat{E}_{sh}^j}{\hat{E}_s^j} = \alpha(W_h^j - W_s^j) + x_{sh}\beta + I(j=3) \cdot x_{sh}\beta^3 + \eta_{sh}^j$$

where $\beta^3 = -\alpha\theta^3$; $I(A)$ is the indicator function such that $I(A)=1$ if A is true and $I(A)=0$ otherwise; hat notation denotes statistical averages; $\eta_{sh}^j = \ln(\hat{E}_{sh}^j / \hat{E}_s^j) - \ln(E_{sh}^j / E_s^j)$ is an error term reflecting sampling error; and we have assumed that $-\alpha f_{sh} = x_{sh}\beta$. The empirical selection and sorting equations are given by

$$(13) \quad \ln \frac{\hat{E}_{sh}^3}{\hat{E}_{sh}^1} - \ln \frac{\hat{E}_s^3}{\hat{E}_s^1} = \alpha[(W_h^3 - W_h^1) - (W_s^3 - W_s^1)] + x_{sh}\gamma + \eta'_{sh},$$

$$(14) \quad \ln \frac{\hat{E}_{sh}^3}{\hat{E}_{sh}^1} = \alpha(W_h^3 - W_h^1) + x_{sh}\gamma + \tau_s + \eta_{sh},$$

where $\gamma = -\alpha(\theta^3 - \theta^1)$, $\eta'_{sh} = \eta_{sh}^3 - \eta_{sh}^1$, and $\eta_{sh} = \ln(\hat{E}_{sh}^3 / \hat{E}_{sh}^1) - \ln(E_{sh}^3 / E_{sh}^1)$.

The key hypothesis is that $\alpha > 0$. If the models are properly specified, all three equations should yield similar estimates of α . To estimate the scale equation (12) we must assume fixed costs are a function of observable characteristics. If that assumption fails, the scale equation may be misspecified. Fixed costs are differenced out of the selection and scale equations, so they should provide a more robust basis for inference. The scale and selection equations require data on both source and destination wages.

This limits the sample, since reliable wage data are not available for all source countries. The sorting equation requires only destination wage data, increasing the number of countries that can be used in the estimation.

For comparison, we also estimate the log-utility model. In the important special case where fixed costs are zero and skill-varying costs are proportional to wages, such that $\lambda m_{sh}^j = -\lambda g_{sh}^j / W_h^j = -\lambda \pi_{sh}$, the empirical counterparts of (8), (9), and (10) are

$$(15) \quad \ln \frac{\hat{E}_{sh}^j}{\hat{E}_s^j} = \lambda (\ln W_h^j - \ln W_s^j) + x_{sh} \theta + \eta_{sh}^j,$$

$$(16) \quad \ln \frac{\hat{E}_{sh}^3}{\hat{E}_{sh}^1} - \ln \frac{\hat{E}_s^3}{\hat{E}_s^1} = \lambda (\delta_h^3 - \delta_s^3) + \eta'_{sh},$$

$$(17) \quad \ln \frac{\hat{E}_{sh}^3}{\hat{E}_{sh}^1} = \lambda \delta_h^3 + \rho_s + \eta_{sh},$$

where we assume that $-\lambda \pi_{sh} = x_{sh} \theta$. The test for income maximization is that $\lambda > 0$. If the models are properly specified, all three equations should yield similar estimates of λ .

3. Data and Empirical Setting

We base our regression analysis on data from Beine, Docquier, and Rapoport (2007; hereafter, BDR), which update the data from Docquier and Marfouk (2006) discussed in section 1. BDR tabulate data on stocks of emigrants by source and destination country. In collaboration with the national statistical offices of OECD countries, they estimate the population in each OECD member of immigrants 25 years and older by source country and education level. In some OECD destinations, these counts are based on census data, whereas in others they are based on register data. BDR classify schooling levels into three categories: primary (0-8 years), secondary (9-12

years), and tertiary (13 plus years). Because education systems differ so much among countries, it is nearly impossible to categorize schooling at a finer level of detail.¹⁰

A. Measurement of Emigrant Stocks

Aggregating data from multiple destination countries raises comparability issues. Some countries, such as Germany, define immigrants on the basis of country of citizenship rather than country of birth. This causes some of the foreign born to be excluded from BDR's immigrant counts in these countries. We check the robustness of our regression results by dropping such countries from some of the specifications. In Belgium and Italy, the statistical office reports aggregate immigrant counts but does not disaggregate by education. BDR impute the skill distribution of immigrants in such cases using data from household surveys. In light of the role that education plays in our analysis, we drop Belgium and Italy from the sample of destinations.

National statistical offices differ in how they classify educational attainment. Some countries' classification systems have no attainment category that distinguishes whether a person who lacks a secondary-school qualification (such as a high school diploma) acquired any secondary education, or whether their schooling stopped at the primary level (grade 8 or below). This could result in inconsistencies in the share of primary-educated immigrants across destination countries. In our regressions we control for whether the destination country explicitly codes primary education.

Some immigrants may have acquired their tertiary schooling in the destination country. By implication, they might have obtained less schooling had they not migrated.

¹⁰ To corroborate BDR, we compared their immigrant stocks for the US with values from Ipums. The two series have correlations of 0.97 (the series will not be exactly identical since Ipums is based on a subsample of the US census and BDR obtained data from special runs conducted by the US Bureau of the Census).

BDR provide some evidence on this point in the form of immigrant counts (for those with tertiary education) that vary by the age at which migrants arrived in the destination country (any age, 12 years or older, 18 years or older, 22 years or older). They find that 68% of tertiary migrants arrive in the destination country at age 22 or older, and 10% arrive between ages 18 and 21, suggesting the large majority of tertiary emigrants depart sending countries at an age at which they would typically have acquired at least some post-secondary education. Reassuringly, the correlations in emigration rates by age at migration range from 0.97 to 0.99. In section 4.2 we provide additional checks on the importance of tertiary schooling acquired in the destination country.

Although our theoretical framework treats migration as a permanent decision, many migrants do not remain abroad forever. There is considerable back-and-forth migration between neighboring countries (Durand, Massey, and Zenteno, 2001), which we address by controlling for source-destination proximity. Furthermore, some migrants are students who will return to their home countries after completing their education. These migrants may have been motivated by educational opportunities in destination countries, as well as wage differences (Rosenzweig, 2006). BDR partially address this issue by restricting the foreign born to be 25 years and older, a population that should have largely completed its schooling. In section 4.2 we attempt to control for differences in educational opportunities between source and destination countries.

Tables 1 and 2 describe broad patterns of migration into OECD countries. As noted in section 1, Table 1 shows that North America receives disproportionately high-skilled migrants, whereas Europe's immigrants are disproportionately low-skilled. Table 2 shows the share of OECD immigrants by country of origin for the 15 largest source

countries. Source countries tend to send emigrants to nearby destinations, as is evident in Turkish migration to Europe, Korean migration to Australia and Oceania, and Mexican and Cuban migration to the United States. Yet, most of the source countries in Table 2 send migrants to all three destination regions.

B. Wage Measures

The key explanatory variables in our regression models are functions of skill-group-specific wages in the source and destination countries.¹¹ Ideally, we would estimate wages by broad education category from the same sources used by BDR. Since such data are not available to us, we turn to different sources.

Our first source is the Luxembourg Income Study (LIS, various years), which collects microdata from the household surveys of 30 primarily developed countries worldwide. This includes most of the destination countries in the BDR data, with the exceptions of Finland, Greece, New Zealand, and Portugal. The intersection of the 13 countries for which BDR and LIS provide useful data (Australia, Austria, Canada, Denmark, France, Germany, Ireland, the Netherlands, Norway, Spain, Sweden, the UK, the US) were host to 91 percent of immigrants in the OECD in 2000.¹² We use data from waves 4 and 5 of the LIS, which span the years 1994-2000.

Although the LIS attempts to “harmonize” the data from different countries, a number of comparability issues arise. One limitation is that the LIS’s constituent household surveys sometimes classify educational attainment differently than the national statistical office of the corresponding country. This adds the problem of within-country

¹¹ An appendix provides more detail on the construction of all wage measures. The data are available for download at <http://harrisschool.uchicago.edu/faculty/web-pages/jeffrey-grogger-migration-files.asp>.

¹² We exclude Switzerland from the destinations because the LIS provides no data on the country after 1992. In 2000, Switzerland had 2.5 percent of the foreign-born population residing in OECD countries.

comparability to the already difficult problem of between-country comparability.

Ultimately, it proved impossible for us to map education categories between the BDR and the LIS data in a manner in which we had full confidence.

Therefore, instead of using education-specific earnings to measure skill-related wages, we use quantiles of each country's earnings distribution. We use the 20th percentile as our measure of low-skill wages and the 80th percentile as our measure of high-skill wages.¹³ We average across 1994 to 2000 for each country in the LIS.

Although the cross-country comparability of the LIS is a desirable feature, we can only use the LIS to estimate our sorting regressions. The reason is that it provides wage data only for our destination countries, whereas the scale and sorting regressions require comparable wage data for the source countries as well. To the best of our knowledge, there is no study that provides micro-level wage data for the 100+ source countries we examine. We rely on two sources of aggregate data to construct the source-destination wage difference measures needed to estimate the scale and sorting regressions.

One combines Gini coefficients from the WIDER World Income Inequality Database with per capita GDP from the World Development Indicators (WDI). Under the assumption that income has a log normal distribution, Gini coefficients can be used to estimate the variance of log income. Using per capita GDP to measure mean annual income, we can then construct estimates of the 20th and 80th percentiles of income, which we are able to do for 102 source and 15 destination countries.

A second source uses data from Freeman and Oostendorp (2000; hereafter FO), who have collected information on earnings by occupation and industry from the

¹³ In a previous version of the paper, we experimented with alternative wage variables based on various measures of low-skill wages and different measures of the return to skill (the standard deviation of income, the Gini coefficient). These alternatives generated results similar to those we report in this paper.

International Labor Organization's October Inquiry Survey. FO standardize the ILO data to correct for differences in how countries report earnings. The resulting data contain observations on earnings in up to 163 occupation-industries per country in each year, from which FO construct deciles for earnings by country and year. For each country, we take as low-skill wages earnings corresponding to the 10th percentile and as high-skill wages earnings corresponding to the 80th percentile. These deciles give the highest correlations with 80th and 20th percentile wages in the LIS. Since not all countries report data in all years, for each country we take the mean across the period 1988 to 1997, creating a sample with 101 source countries and 12 destinations.

Table 3 presents summary statistics of these wage measures. The top two panels provide data for the destination countries. The top panel shows that the LIS produces higher wages and larger skill-related wage differences than the other sources. Despite the differences in scale, the correlation between skill-related wage differences in the LIS and the WDI data is 0.86; between the LIS and the FO data it is 0.78.¹⁴

The second panel reports summary statistics for after-tax measures of destination-country wages. We consider such measures since pre-tax wage differences overstate the return to skill enjoyed by workers and since tax policy varies within the OECD (Alesina and Angeletos, 2002). To construct post-tax wage differences we employ average tax rates by income level published by the OECD since 1996 (OECD, various years).¹⁵ To

¹⁴ Using household level data from the World Bank, Clemons, Montenegro, and Pritchett (2008) estimate the ratio of US wages to national wages by education level (primary, secondary, tertiary) for 32 of the developing countries in our sample. Skill related income differences (relative to the US) in their data and in our data have a correlation of 0.55 for WDI wages (where we match wages at the 20th (80th) percentile in our data to wages for individuals with primary (tertiary) education in their data) and 0.44 for FO wages. To control for differences in scaling, we force US wages by skill level to be the same in their data and in ours.

¹⁵ Prior to averaging income across years, we match to each year and income group that year's corresponding tax rate. Since the tax data only go back to 1996, we use tax rates for that year to calculate post-tax income values in 1994 and 1995.

20th percentile earnings we apply the tax rate applicable to single workers with no dependents whose earnings equal 67 percent of the average production worker's earnings. To 80th percentile earnings we apply the tax rate applicable to a comparable worker with earnings equal to 167 percent of the average production worker's earnings. After-tax wage differences are only about half as large as pre-tax differences.¹⁶

The third panel provides data for the source countries. Only WDI and FO data are shown, since the LIS provides no source-country data. Source country wages vary less than destination-country wages between the two sources; the correlation between skill-related wage differences is 0.91. Unfortunately, we have no tax data for most of our source countries. Thus the scale and selection regressions below are estimated only from pre-tax wage data, whereas we report sorting regressions for pre- and post-tax wages.

D. Other Variables in the Regression Model

Differences in language between source and destination countries may be relatively more important for more-educated workers, since communication and information processing are likely to be salient aspects of their occupations. We control for whether the source and destination country share a common official language based on data from CEPII (<http://www.cepii.fr/>). Similarly, English-speaking countries may attract skilled emigrants because English is widely taught in school as a second language.¹⁷ To avoid confounding destination-country skilled-unskilled wage differences

¹⁶ We assume natives and immigrants face the same tax rates. Recently, some countries have begun to offer tax breaks to high skilled immigrants to encourage skilled immigration (e.g., the Netherlands). This practice appears to have been uncommon during the 1990s, which is the period for our income data.

¹⁷ English-speaking countries may also attract the more skilled because they have legal traditions based on common law that provide relatively strong protection of property rights (Glaeser and Schleifer, 2002).

with the attraction of being in an English-speaking country, we control for whether a destination country has English as its primary language.

Migration costs are likely to be increasing in distance between countries. Proximity may make illegal immigration less costly, thereby increasing the relative migration of the less-skilled. We include as regressors great circle distance, the absolute difference in longitude, and an indicator for source-destination contiguity. Migration networks may lower migration costs (Munshi, 2003), possibly benefiting lower-income individuals more (Orrenius and Zavodny, 2005; McKenzie and Rapoport, 2006). As networks may be stronger between countries that share a common colonial heritage, we include CEPII's indicators of whether a pair of countries have short or long colonial histories. We also control for migrant networks using lagged migration, measured as the total stock of emigrants from a source country in a destination as of 1990.¹⁸

Destination countries impose conditions in deciding which immigrants to admit, many of which involve education. One indicator of skill bias in a country's admission policies is the fraction of visas it reserves for refugees and asylees. Less-educated individuals may be more likely to end up as refugees, making countries that favor refugees in their admissions likely to receive more less-educated immigrants. We control for the share of immigrant inflows composed of refugees and asylees averaged over the 1992-1999 period (OECD, 2005).¹⁹ European signatories of the Schengen Agreement have committed to abolish all border barriers, including temporary migration restrictions,

¹⁸ Because we are missing lagged migration for many observations in the sample, we add the variable only in later specifications. All results are robust to its inclusion.

¹⁹ Countries also differ in the share of visas that they reserve for skilled labor. Unfortunately, we could only obtain this measure for a subset of destination countries. Over time, the share of visas awarded to asylees/refugees and the share awarded to skill workers are strongly negative correlated (OECD, 2005), suggesting policies on asylees/refugees may be a sufficient statistic for a country's immigration priorities.

on participating countries. We control for whether a source-destination pair were both signatories of Schengen as of 1999. Similarly, some countries do not require visas for particular source countries, with the set of visa-waiver countries varying across destinations. While visa waivers strictly affect only tourist and business travelers, they may indicate a source-country bias that also applies to other immigrant admissions. We control for whether a destination country grants a visa waiver to individuals from a source country as of 1999. Clearly, other aspects of policy may influence migration as well. Unfortunately, existing data give little indication of how immigration policies vary across destinations. As important as immigration policy may be, existing data simply do not permit a more detailed characterization of the policy environment.

Finally, note that the regressors used in the analysis vary either by destination or source-destination pair. Other source-country-specific characteristics could also affect migration. Some, such as the state of the credit market or the poverty rate, are observable and could be controlled for explicitly. Others, however, are unobservable. Rather than controlling for a limited set of observable source-country characteristics explicitly, we provide implicit controls for both observable and unobservable source-country characteristics via the source-country fixed effects in the sorting regression.²⁰

4. Regression Analysis

A. Main results

Our main regression analyses are based on the scale, selection, and sorting regressions derived from the linear-utility model, as shown in equations (12), (13), and

²⁰ Docquier, Lohest, and Marfouk (2007) examine the source-country correlates of the scale of migration. In unreported results, we experimented with two source-specific variables. Private credit to the private sector as a share of GDP is a measure of the financial development of the source country (Aghion et al. 2006), which may affect constraints on financing migration. The incidence of poverty in the source country, as proxied by the share of agriculture in GDP (Belot and Hatton, 2008), may also affect credit constraints. The inclusion of neither variable affected our core results.

(14), respectively. Our main results are based on wage measures constructed from the WDI and LIS data. Estimates are reported in Table 4.

In the scale equation reported in column (1), the unit of observation is the source-destination-skill group cell, with one observation for the primary educated ($j=1$) and one observation for the tertiary educated ($j=3$) for each source-destination pair. The dependent variable is the log odds of emigrating from source s to destination h for members of skill group j , and the wage measure is the skill-specific difference in pre-tax wages between the destination and source countries, $W_h^j - W_s^j$. In the selection equation reported in column (2), the unit of observation is the source-destination pair.²¹ The dependent variable is the difference between the log skill ratio of emigrants from s to h and the log skill ratio of non-migrants in source s .²² The wage measure is the difference between the destination and the source in skill-related pre-tax wage differences, $(W_h^3 - W_h^1) - (W_s^3 - W_s^1)$. In the sorting equations reported in columns (3) through (6), the unit of observation is again the source-destination pair, but the dependent variable is the log skill ratio of emigrants from s to h . The key independent variable is the skill-related wage difference of the destination country, $(W_h^3 - W_h^1)$. Like the scale and selection regressions, the sorting regressions in columns (3) and (4) are based on the WDI data; column (3) is based on pre-tax data, whereas column (4) is based on post-tax data. Columns (5) and (6) are based on pre- and post-tax data from the LIS.

²¹ In the WDI data, there are 15 destinations and 102 source countries. Since source countries do not send emigrants to every destination country, the number of observations is less than $15 \times 102 = 1530$.

²² Equivalently, the dependent variable can be seen as the difference in the log odds of migrating from source s to destination h between the tertiary educated and the primary educated.

Because the dependent variables have a log-odds metric, the magnitude of the regression coefficients does not have a particularly useful interpretation.²³ We focus in this section on the signs and significance levels of the coefficients. Below, we discuss applications that provide information about the quantitative effects of key variables. In addition to the variables shown, all of the regressions include a dummy variable equal to one if the destination-country statistical office explicitly codes a primary education category. This controls for systematic differences in our dependent variable that arise from different coding schemes, as discussed in section 3. The scale regression includes a dummy variable equal to one for observations corresponding to the tertiary-educated skill group, denoted $I(j=3)$, and interactions between that dummy and all other regressors (not shown to save space). The sorting regressions include a full set of source-country dummies. Standard errors, reported in parentheses, are clustered by destination country.

The wage coefficients in columns (1) through (3) are directly comparable because they are all based on pre-tax data from WDI. In the context of our model, they each provide estimates of the same parameter α , where income maximization implies $\alpha > 0$. Furthermore, if the regression models are properly specified, the coefficients from scale, selection, and sorting regressions should be the same.

In Table 4, all three wage coefficients are positive, as predicted by theory. The coefficients in the selection and sorting regressions are similar in value and statistically significant. However, the coefficient in the scale equation is smaller and insignificant. This may indicate that omitted fixed costs result in a misspecified scale equation. In the scale equation we assume fixed costs are a function of observable characteristics of the

²³ Based on (4), one might think the coefficient on earnings would identify the marginal utility of income. However, this would only be true if the variance on the idiosyncratic component of utility in (3) is unity.

source-destination pair, while the selection and sorting regressions difference out fixed costs. The difference in wage coefficients between the scale and selection regressions suggests the scale equation omits fixed costs that are negatively correlated with the difference in skill-specific wage differences between destination and source countries.

The wage coefficient in column (4) suggests that migrants sort more strongly on post-tax wages than pre-tax wages, as one might expect. The estimates in columns (5) and (6), based on wage data from the LIS, show a similar pattern. Both coefficients are positive and significant, and the coefficient on post-tax wages in column (6) is larger than the coefficient on pre-tax wages in column (5). Among the destination countries in the sample, the U.S and Canada have relatively large pre-tax skill-related wage differences. Since these countries also have less progressive tax systems, their relative attractiveness to skilled migrants is enhanced by accounting for taxes.

The regressions also include variables reflecting geographic, linguistic, social, and political relationships between countries. The positive coefficient on the Anglophone-destination dummy in column (1) shows that English-speaking countries receive more immigrants than other countries, all else equal. The coefficient in the selection regression (column (2)) shows that emigrants bound for English-speaking destinations are more highly educated in relation to their non-migrant countrymen than emigrants bound elsewhere. The coefficients in the sorting regressions (columns (3) through (6)) show that English-speaking destinations tend to attract higher-skilled immigrants. Emigration is also greater toward destinations that share a common language with the source, and such emigrants are more skilled than either their non-migrant counterparts or emigrants

from the same source bound to other destinations. This suggests that migrants perceive higher rewards to skill in destinations where they can speak a language they know.²⁴

The next three variables capture the effects of geography. Contiguity raises the scale of migration. However, it reduces the skills of emigrants, all else equal, in relation both to non-migrants (as seen in the selection regression) and to migrants to non-contiguous destinations (as seen in the sorting regression), perhaps reflecting the ease of illegal migration between neighboring countries. In the scale equation, the longitude-difference coefficient is insignificant, but the log-distance coefficient is negative and significant. One interpretation is that migration is lower, the greater the distance between the source and the destination, but controlling for distance, the need to cross an ocean (which follows from large longitudinal distances) has no independent effect.

Colonial relationships increase the scale of migration, all else equal. At the same time, emigrants to the former colonial power are less skilled than non-migrants and less skilled than emigrants to other destinations. Recent literature suggests that economic and social networks between industrialized countries and their former colonies contribute to bilateral migration flows, much in the way such networks also appear to contribute to bilateral trade (Pedersen, Pytlikova, and Smith, 2004; Mayda, 2005). Our results are consistent with these linkages disproportionately affecting migration of the less-skilled.

There is also an important role for our measures of immigration policy.²⁵ The effect of asylum policy on the scale of immigration is insignificant, but generous asylum policies reduce immigrant skills with relation to both non-migrants and migrants to other destinations. This finding suggests destinations that allocate a higher share of visas to

²⁴ These results and those for distance and colonial history are consistent with Belot and Hatton (2008).

²⁵ For work on migration policy impacts, see Mayda (2005) and Clark, Hatton, and Williamson (2007).

asylees and refugees may limit opportunities for more-skilled migrants to gain entry, producing a less skilled migrant inflow.²⁶ Visa waivers are associated with higher migration rates, although the effect is marginally significant. Visa waivers significantly reduce the skills of emigrants in relation to non-migrants, but increase skill in relation to emigrants who move to a destination with which the source country has no visa waiver. The Schengen accord has had little effect on the scale of migration among signatory countries, but it is associated with positive selection and positive sorting of migrants.

B. Results for the Log Utility Model

Table 5 reports results based on the scale, selection, and sorting regressions derived from the log-utility model in equations (15), (16), and (17). The dependent variables in Table 5 are the same as those in the corresponding columns of Table 4. The wage measures differ between the linear and log-utility models. In the scale equation of the log-utility model, reported in column (1), the wage measure is the skill-specific difference in pre-tax log wages between the destination and source countries, $\ln W_h^j - \ln W_s^j$. In the selection equation reported in column (2), the measure is the difference between the destination and the source in the return to skill, $(\delta_h^3 - \delta_s^1)$, where the return to skill in a country is the log ratio of high-skill to low-skill wages. In the sorting equations reported in columns (3) through (6), the wage variable is the return to skill in the destination country, δ_h^3 . Columns (1)-(4) are based on the WDI data, whereas columns (5) and (6) are based on LIS data. Returns to skill are based on pre-tax data in columns (1)-(3), and (5) and on post-tax data in columns (4) and (6). To focus on a case

²⁶ On asylee and refugee policy in Europe, see Hatton and Williamson (2004).

of special importance in the literature, we impose the assumptions that fixed migration costs are zero and skill-varying costs are proportionate to wages. This implies that in the scale regression the regressors control for proportional migration costs (see (15) and surrounding discussion). It also means that the only regressor in the selection regression is $(\delta_h^3 - \delta_s^1)$, since proportional costs are differenced out. Likewise it implies that the only regressors in the sorting regressions are δ_h^3 and the source-country dummies.

As in the linear-utility model, utility maximization implies that all of the coefficients on log wages and returns to skill should be positive. Furthermore, if the model is properly specified, the coefficients in columns (1) through (3) should be similar. In fact, the wage coefficients in the scale and selection regressions are negative and significant, whereas the sorting coefficients are both positive and significant.

The assumptions that fixed costs are zero and skill-varying costs are proportional to wages result in rather sparsely parameterized regressions.²⁷ When we relax these restrictions by assuming both fixed and skill-varying costs to be functions of observed country-pair characteristics, the wage coefficients in the scale and selection regressions remain negative and significant and the wage coefficients in the sorting regressions remain positive and significant.²⁸ Thus, the sign pattern of the coefficients in Table 5 holds whether or not other regressors are included in the estimation.

²⁷ Belot and Hatton (2008) find that the correlation between skilled migration rates and the skill-specific difference in log wages between source and destination countries is sensitive to whether controls for poverty rates in the source are included in the estimation. In unreported results, we find that the negative coefficient on the returns to skill we estimate in the log utility selection regression obtains whether or not controls for poverty rates are included in the estimation (see note 20).

²⁸ In the log utility model, if we assume that fixed migration costs are a function of the same variables as in Table 4, allowing for fixed costs means including these variables as regressors, divided by the destination country wage, as shown in the derivations of equations (8)-(10). Alternatively, one might imagine including these regressors uninteracted with the destination wage. Under either specification, the log wage variable enters with a negative sign in the scale and selection regressions.

We see two potential explanations for the difference between the linear-utility and log-utility regressions. One concerns omitted variable bias due to weak controls for fixed costs. Differencing the scale equation between skill groups eliminates fixed costs from the selection and sorting regressions in the case of linear utility, but not in the case of log utility. Fixed costs that were strongly negatively correlated with source-destination differences in log wages and returns to skill could explain the negative coefficients in the scale and selection regressions in Table 5.

Perhaps more important is the lack of negative selectivity in the data, as seen in Figure 1. Log-utility maximization requires $\lambda > 0$. It also requires that for destination-source pairs where $\delta_h^3 - \delta_s^3 < 0$, migrants be negatively selected. In the data, there are many cases where $\delta_h^3 - \delta_s^3 < 0$, but no negative selection, as seen by comparing Figures 2 and 3. Inspection of equation (9) shows that such negative correlation between $\delta_h^3 - \delta_s^3$ and $\ln(E_{sh}^3 / E_{sh}^1) - \ln(E_s^3 / E_s^1)$ will tend to result in a negative estimate of λ , contrary to the requirements of the theory. In other words, the lack of negative selection in the data is at odds with the assumption that migrants maximize the log utility of net wages.²⁹

A remaining question is why the wage coefficients in the log-utility sorting regressions are positive, like their counterparts in the linear-utility sorting regressions. Why do the sorting regressions fail to distinguish between linear and log utility, when the selection regressions draw the distinction so clearly? The reason is that the wage measure only varies among the 15 destination countries, and among countries with

²⁹ Strictly speaking, it is at odds with the joint assumptions that migrants maximize log utility and the extreme value disturbances are multiplicative. Since we continue to estimate a negative coefficient on log wages when other regressors are added to the selection regression, the assumption of proportional migration costs does not appear to account for the failure of the log utility model.

similar labor productivity sorting on log differences in wages looks similar to sorting on level differences in wages. The rank correlation between the log wage difference and the level wage difference across destinations is 0.68. In order to distinguish between linear and log utility on the basis of the sorting regressions, one would need a sample that included destinations with more widely differing levels of productivity.³⁰

C. Robustness Checks

Tables 6-8 report specifications designed to check the robustness of our results. We restrict attention to the linear utility model in light of its superior performance relative to the log utility model. We further restrict attention to the selection and sorting regressions, since they are more robust in the presence of fixed migration costs. For the sorting regressions, we focus on specifications with post-tax wage differences. All regressions include the variables in our baseline specifications, shown in columns (2), (4) and (6) of Table 4. We present only the wage coefficients in order to conserve space.

Table 6 presents estimates based on alternative wage measures. The top panel reports results based on WDI wages in which source wages are adjusted by source-country PPP and destination wages are adjusted by destination PPP, to account for differences in the cost of living across countries. Adjusting for PPP makes the coefficient in the selection regression slightly larger and the coefficient in the sorting regression slightly smaller and insignificant. In the second panel, we see that adjusting for PPP using LIS wages yields wage coefficients that are positive and significant, as in Table 4.

The bottom two panels of Table 6 present results based on the Freeman-Oostendorp wage data described in Section 3. Without adjusting for PPP, the wage

³⁰ The similarity of productivity levels among US states may explain why log-utility models have yielded evidence in favor of sorting among US domestic migrants (Borjas, Bronars, and Trejo 1994; Dahl 2002).

coefficients in both the selection and sorting regressions are positive. The selection coefficient is significant, whereas the sorting coefficient has a t-statistic of 1.6. Adjusting the Freeman-Oostendorp wages for PPP reduces the selection coefficient and raises both the sorting coefficient and its significance. The key results from the linear utility model thus appear to be fairly robust to alternative wage measures.

In Table 7 we return to our original unadjusted, WDI and LIS-based wage measures and report results obtained from alternative specifications. Columns (1) through (3) address the problem that some emigrants may have obtained their tertiary education in the destination country rather than the source country. If the cost of acquiring tertiary education across destination countries were negatively correlated with destination-country wage differences, then the effect on immigrant skill that we attribute to wage differences could instead be due to differences in educational costs. To deal with this issue we redefine the numerator of the skill ratios in the dependent variables to be the sum of tertiary- and secondary-educated immigrants. This addresses the problem if we can assume all tertiary-educated immigrants would have obtained at least a secondary education in their source country. The coefficients in columns (1) through (3), where the dependent variables are based on this alternative definition of the log skill ratio, are all positive and significant and differ little from estimates in our baseline specifications.

Columns (4) through (9) report the results of adding to our baseline specifications two variables designed to capture other potential costs or benefits of migration that vary by skill. Columns (4)-(6) add a relative university quality measure based on the world-wide ranking of universities by Shanghai Jiao Tong University (<http://ed.sjtu.edu.cn>). It is equal to the average rank of universities within the destination country (among top 250

universities worldwide), interacted with a dummy variable equal to one if the source country has no ranked universities. We intend this as a proxy for the education-related benefit of migrating relative to remaining in the home country. Relative university quality has no effect on emigrant selectivity, as seen in column (4). The coefficients in the sorting regressions (columns (5) and (6)) are negative, as one might expect (higher-ranked institutions have ranks closer to one), and significant. Higher ranked universities appear to act as a draw for higher-skilled immigrants from countries with low-quality education systems, consistent with Rosenzweig (2006). The wage coefficients in all three regressions are similar to those from our baseline specifications.

Columns (7) through (9) add the log total stock of emigrants from the source in the destination as of 1990. We are missing this variable for about 30% of our sample, which causes the number of observations to drop. Nevertheless, the wage variables have similar magnitudes and patterns of significance as in Table 4. In the selection regression, the lagged migrant stock enters with a negative sign and is precisely estimated. Larger past bilateral migration is associated with less-educated current migration, consistent with migrant networks lowering migration costs disproportionately for the less-skilled (Orrenius and Zavodny, 2005; McKenzie and Rapoport, 2006; Clark, Hatton and Williamson, 2007). In the sorting regressions, lagged migration also enters negatively, indicating that the pull of an existing migrant stock in a destination is stronger for less-skilled migrants, but the coefficient is precisely estimated only in one of the regressions.

In columns (10) and (11), we present sorting regressions based on data from all the available source countries, irrespective of whether we have wage data for them. This

highlights the advantage of the sorting regression, for which only destination-country wage data is necessary. The results are similar to those for the smaller sample.

Table 8 addresses the independence of irrelevant alternatives (IIA) assumption implicit in the conditional logit framework. IIA arises from the assumption that the error terms in equation (3) are i.i.d. across alternative destinations. IIA may be violated if two or more of our destinations are perceived as close substitutes by potential migrants. Hausman and McFadden (1984) note that if IIA is satisfied, then the estimated regression coefficients should be stable across choice sets. In the context of our application, this means that the regression coefficients should be similar when we drop destinations from the sample. To check for violations of IIA, we re-estimated our models 15 times, each time dropping one of the 15 destinations. The resulting coefficients on the key wage variables are reported in Table 8. In general, they are quite similar across samples, suggesting that the IIA property is not violated in our data.³¹

D. Fixed Costs and the Sorting of Migrants by Skill Level

In this section we use our parameter estimates to shed light on different dimensions of international migration. The first issue we address concerns fixed costs, which play a role in determining the scale of migration. Our framework allows us to estimate migration costs that are specific to each source-destination pair.

The estimates stem from the scale equation (4). If we include a dummy variable for each source-destination pair in our sample, assuming as before that skill-varying costs are given by $g_{sh}^j = x_{sh}\theta^j$, we obtain numerically identical estimates to those obtained by

³¹ We attempted to compute asymptotic chi-square statistics along the lines of Hausman and McFadden (1984) to test for stability across choice sets in all the regression coefficients. For the most part, the asymptotic covariance matrices were singular, a finite-sample problem that often arises in Hausman tests.

estimating the selection equation (13). However, as a by-product, we obtain estimates of $-\alpha f_{sh}$ from the coefficients on the source-destination dummy variables. To recover an estimate of fixed costs f_{sh} , we divide those coefficients by our estimate of $-\alpha$, where α is the coefficient on wages. This provides estimates of fixed migration costs relative to an omitted source-destination base pair, in thousands of 2000 US dollars per year (the units in which wages are measured). We choose the Mexico-US pair as the base since it involves the largest migration flow. Of course, these estimates reflect not only direct monetary costs, but also the monetary value of psychic costs and source-specific immigration policies imposed by the destination countries.

Estimates for each source-destination pair in our sample are shown in an online appendix.³² Table 9 presents estimates for the subset of source and destination countries that appear in the 25 source-destination pairs with the largest stocks of migrants. Within each source-destination cell, the first entry is the estimated fixed migration cost. The second entry is the number of emigrants from the source to the destination.

The US is the low-cost destination for all the Western Hemisphere source countries except Jamaica, and it receives more emigrants from those countries than any other destination. At the same time, migration costs are only part of the story. For Chinese emigrants, the cost of migrating to Canada and the US is about the same. Yet many more go the US, presumably due to the higher wages there. The situation is similar for German emigrants. Canada, France, and the UK are all lower-cost destinations than the US, yet the US has more German immigrants than those three destinations combined.

Several entries highlight the role of history. Germany is by far the lowest-cost destination for Turkish emigrants, despite Turkey's similar proximity to the other

³² Available at <http://harrisschool.uchicago.edu/faculty/web-pages/jeffrey-grogger-migration-files.asp>.

European countries. The US is the low-cost destination for Vietnam, despite the country's proximity to Australia and colonial ties to France. Presumably, these estimates reflect Germany's labor-recruitment strategy from the 1960s, America's post-war asylum policy in the 1980s, and the immigrant networks that have developed in their wake.

We next ask how wage differences and skill-varying migration costs explain differences in mean immigrant skills among the destination countries. The first column of Table 10 presents our measure of immigrant skills, which is the mean log skill ratio among immigrants in each destination country. Based on this measure, the US has the most highly skilled immigrants on average, followed by Ireland and Canada.³³ We seek to explain the immigrant skill gap, defined as the difference between the mean log skill ratio among immigrants in the US and the mean log skill ratio among immigrants in other destination countries. The immigrant skill gap is reported in column (2).

We use the sorting regression reported in column (4) of Table 4 to carry out the decomposition, which explains the immigrant skills gap as a linear combination of the differences in mean values of the regressors, using the regression coefficients as weights. To aid interpretation, we report results in the form of the share of the immigrant skill gap explained by each variable in the regression.³⁴

Results are reported in columns (3) through (14). Column (3) shows that on average the wage difference explains 58 percent of the immigrant skill gap; in all destination countries it explains at least 25 percent. In Ireland, which has a relatively small immigrant skill gap, it explains over 100 percent. The next two columns show the

³³ Other skill measures give somewhat different rankings. For example, Canada ranks first in the share of immigrants with tertiary education. The reason for the difference is that the US has a lower share of primary-educated immigrants than Canada. We focus on the log skill ratio because that is the skill measure that follows from our model and the measure for which our regressions can provide a decomposition.

³⁴ Nothing constrains the share explained by any subset of components to be less than one.

importance of language. English explains at least 20 percent of the immigrant skill gap for each non-Anglophone destination country. The role of common languages is smaller overall, but nevertheless important for some of those destination countries whose languages are not widely spoken elsewhere. Among the policy variables, visa waivers and the Schengen treaty explain relatively little of the immigrant wage gap. Asylum policy, in contrast, has important effects. In seven of the destination countries, asylum policy explains at least 20 percent of the immigrant skill disadvantage. In Canada and New Zealand, in contrast, the skills gap would be over 10 percent larger were it not for their relatively restrictive admissions of asylum seekers.

5. Conclusions

Two dominant features of international labor movements are positive selection of individuals into migration and positive sorting of migrants across destinations. We show that a simple model of income maximization can account for both phenomena.

In our selection regression, we find that migrants for a source-destination pair are more educated relative to non-migrants, the larger is the skill-related difference in earnings between the destination country and the source. That is, positive selectivity is stronger where the reward to skill in the destination is relatively large. This result obtains for wage differences expressed in levels, but not in logs. Log wage differences, which capture cross-country differences in returns to skill, fail to account for bilateral migration patterns because cross-country differences in returns to skill are dwarfed by cross-country differences in labor productivity. On their own, cross-country differences in returns to skill would predict negative selection of migrants, which occurs rarely in the data.

Positive sorting is a general prediction of income maximization. In our sorting regression, the relative stock of more-educated migrants in a destination is increasing in the level earnings difference between high and low-skilled workers. This correlation is stronger when wage differences are adjusted for taxes, implying that migrants weigh post-tax earnings when choosing a destination. The US and Canada enjoy relatively large post-tax skill-related wage differences, which largely account for their ability to attract more educated migrants relative to other OECD countries.

In the sorting regression, we obtain qualitatively similar results when we use wages constructed from micro data as when we approximate wages using aggregate income data and impose the assumption of log normality. As a practical matter, this means that one can obtain empirically meaningful estimates of skill-related wage differences from commonly available data sources. The sorting regression allows one to test income maximization even without source-country wage data, which makes our approach applicable to a wide variety of settings.

Destinations with liberal refugee and asylum policies draw relatively low-skilled immigrants, all else equal. Unfortunately, our ability to say more about policy is limited by the sparseness of data allowing one to compare the regimes of different destination countries. Our model provides a framework in which comparative analysis of immigration policies could be undertaken, but with current limitations in data we are limited in the analyses we can carry out.

References

- Adams, Richard. 2003. "International Migration, Remittances, and Brain Drain: A Study of 24 Labor Exporting Countries." World Bank Policy Research Working Paper no. 2972.
- Aghion, Philippe, Philippe Bacchetta, Romain Ranciere, and Kenneth Rogoff. 2006. "Exchange Rate Volatility and Productivity Growth: The Role of Financial Development." NBER Working Paper No. 12117.
- Alesina, Alberto, and George-Marios Angeletos. 2003. "Fairness and Redistribution: US versus Europe." NBER Working Paper No. 9502.
- Beine, Michael, Frédéric Docquier, and Hillel Rapoport. 2001. "Brain Drain and Economic Growth: Theory and Evidence." *Journal of Development Economics*, 64(1): 275-289.
- Beine, Michel, Frédéric Docquier, and Hillel Rapoport. 2007. "Measuring International Skilled Migration: New Estimates Accounting for Age of Entry." *World Bank Economic Review*, 21(June 2): 249-254.
- Beine, Michel, Frédéric Docquier, and Hillel Rapoport. 2008. "Brain Drain and Human Capital Formation in Developing Countries: Winners and Losers." *Economic Journal*, forthcoming.
- Belot, and Timothy Hatton. 2008. "Immigrant Selection in the OECD." CEPR Working Paper No. 6675.
- Bendel, R. B., Higgins, S. S., Teberg, J.E., and Pyke, D.A. "Comparison of skewness coefficient, coefficient of variation, and Gini coefficient as inequality measures within populations." *Oecologia* 78 (1989), 394-400.
- Borjas, George J. 1987. "Self-Selection and the Earnings of Immigrants." *American Economic Review* 77(4): 531-553.
- Borjas, George, J. 1991. "Immigration and Self-Selection." In John Abowd and Richard Freeman, eds., *Immigration, Trade, and the Labor Market*. Chicago: University of Chicago Press, pp. 29-76.
- Borjas, George J. 1999. "The Economic Analysis of Immigration." In Orley C. Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Amsterdam: North-Holland, pp. 1697-1760.
- Borjas, George J. 2006. "Labor Outflows and Labor Inflows in Puerto Rico." Mimeo, Harvard University.
- Borjas, George J., Steven G. Bronars, and Stephen Trejo. 1992. "Self-Selection and Internal Migration in the United States." *Journal of Urban Economics* 32: 159-185.
- Borjas, George J., and Lynette Hilton. 1996. "Immigration and the Welfare State: Immigrant Participation in Means-Tested Entitlement Programs" *Quarterly Journal of Economics*, 111(2): 575-604.
- Brücker, Herbert, and Cécily Defoort. 2006. "The (Self) Selection of International Migrants Reconsidered: Theory and New Evidence." IZA Discussion Paper 2052.
- Carrington, William J. and Enrica Detragiache. 1998. "How Big is the Brain Drain?" IMF Working Paper 98/102.
- Chiquiar, Daniel and Gordon Hanson. 2005. "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States." *Journal of Political Economy*, April 113(2): 239-281.

- Clark, Ximena, Timothy Hatton, Jeffrey Williamson. 2007. "Explaining US Immigration, 1971-1998." *Review of Economics and Statistics*, 89(2): 359-373.
- Clemons, Michael, Claudio Montenegro, and Lant Pritchett. 2008. "The Great Discrimination: Borders as a Labor Market Barrier." Mimeo, Harvard University.
- Dahl, Gordon. 2002. "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets." *Econometrica* 70(6): 2367-2420.
- Docquier, Frédéric and Abdeslam Marfouk. 2006. "International Migration by Educational Attainment, 1990-2000." In Caglar Ozden and Maurice Schiff, eds., *International Migration, Remittances, and the Brain Drain*. Washington, DC: The World Bank and Palgrave MacMillan, pp. 151-200.
- Docquier, Frédéric, Olivier Lohest, and Abdeslam Marfouk. 2007. "Brain Drain in Developing Countries." *World Bank Economic Review*, 21 (2): 193-218.
- Docquier, Frédéric, and Hillel Rapoport. 2007. "Skilled Migration. The Perspective of Developing Countries."
- Durand, Jorge, Douglas S. Massey, and Rene M. Zenteno. 2001. "Mexican Immigration in the United States." *Latin American Research Review* 36(1): 107-127.
- Fernandez-Huertas, Jesus. 2006. "New Evidence on Emigration Selection." Mimeo, Columbia University.
- Freeman, Richard B. 2006. "People Flows in Globalization." NBER Working Paper No.12315.
- Freeman, Richard B., and Remco H. Oostendorp. 2000. "Wages around the World: Pay across Occupations and Countries." NBER Working Paper No. 8058.
- Glaeser, Edward I. and Andrei Shleifer. "Legal Origins," *Quarterly Journal of Economics*, 2002, v107(4,Nov), 1193-1229.
- Grogger, Jeffrey, and Gordon H. Hanson. 2007. "Income Maximization and the Sorting of Emigrants across Destination Countries." Mimeo, UCSD.
- Hanson, Gordon H. 2007. "International Migration and the Developing World." Mimeo, UCSD.
- Hatton, Timothy J. and Jeffrey G. Williamson. 2004. "Refugees, Asylum Seekers, and Policy in Europe." NBER Working Paper No. 10680
- Hatton, Timothy J. and Jeffrey G. Williamson. 2005. "What Fundamentals Drive World Migration?" In George J. Borjas and Jeff Crisp, eds., *Poverty, International Migration, and Asylum*, New York: MacMillan.
- Hanushek, Eric A., and Lei Zhang. 2006. "Quality-Consistent Estimates of International Returns to Skill." NBER Working Paper No. 12664.
- Hausman, Jerry, and Daniel McFadden. 1984. "Specification Tests for the Multinomial Logit Model." *Econometrica*, 52(5): 1219-1240.
- Ibarraran, Pablo, and Darren Lubotsky. 2005. "The Socioeconomic Status of Mexican Migrant Families: New Evidence from the 2000 Mexican Census." In George J. Borjas, ed., *Mexican Immigration*. Chicago: University of Chicago and National Bureau of Economic Research, forthcoming.
- Johnson, Norman, and Samuel Kotz. *Continuous Univariate Distributions-1*. New York: John Wiley and Sons, 1970.
- Kapur, Devesh, and John McHale. 2005. *Give Us Your Best and Brightest: The Global Hunt for Talent and its Impact on the Developing World*. Washington, DC: Center for Global Development and the Brookings Institution.

- Luxembourg Income Study (LIS) Micro database, (various years); harmonization of original surveys conducted by the Luxembourg Income Study, Asbl.
- Mayda, Anna Maria. 2005. "International Migration: A Panel Data Analysis of Economic and Non-economic Determinants." IZA Discussion Paper No. 1590.
- McFadden, Daniel. 1974. "The Measurement of Urban Travel Demand." *Journal of Public Economics*, 3: 303-328.
- McKenzie, David and Hillel Rapoport. 2006. "Self-Selection Patterns in Mexico-US Migration: The Role of Migration Networks." Mimeo, the World Bank and Bar-Ilan University.
- Munshi, Kaivan. 2003. "Networks in the Modern Economy: Mexican Migrants in the US Labor Market." *Quarterly Journal of Economics* 118 (May): 549-97.
- OECD. 2003. *Database on Foreign-born and Expatriates*. <http://www.oecd.org/>.
- OECD. 2005. *Trends in International Migration*. OECD: Paris.
- OECD. Various years. *Taxing Wages*. OECD: Paris.
- Orrenius, Pia M. and Madeline Zavodny. 2005. "Self-Selection among Undocumented Immigrants from Mexico." *Journal of Development Economics*, October 78(1): 215-240.
- Pedersen, Peder J., Mariola Pytlikova, and Nina Smith. 2004. "Selection or Network Effects? Migration Flows into 27 OECD Countries, 1990-2000," IZA Discussion paper 1104.
- Psacharopoulos George, Harry Anthony Patrinos. 2004. "Returns to Investment in Education: A Further Update." *Economics of Education*, 12(2): 111-134.
- Ramos, Fernando. 1992. "Out-Migration and Return Migration of Puerto Ricans." In *Immigration and the Work Force*, George Borjas and Richard Freeman (eds.). Chicago: University of Chicago Press and NBER. 49-66.
- Rosenzweig, Mark. 2006. "Global Wage Differences and International Student Flows." *Brookings Trade Forum*.
- Rosenzweig, Mark. 2007. "Education and Migration: A Global Perspective." Mimeo, Yale.
- Roy, A.D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3 (June): 135-46.
- United Nations. 2006. "World Migrant Stock: The 2005 Revision Population Database" <http://esa.un.org/migration/>.
- Volger, Michael, and Rotte, Ralph. 2000. "The Effects of Development on Migration: Theoretical Issues and New Empirical Evidence." *Journal of Population Economics*, 13(3): 485-508.
- Yang, Dean. 2006. "Why Do Migrants Return to Poor Countries? Evidence from Philippine Migrants' Responses to Exchange Rate Shocks." *Review of Economics and Statistics*, Vol. 88, No. 4, November 2006, pp. 715-735.

Appendix: Wage Data

LIS wage measures

We use the LIS to measure the 20th and 80th percentiles of the annual wage distribution for destination countries. Our wage measure is gross cash wage and salary income. To limit quantity variation in measuring the price of labor, we restrict the samples to male household heads between the ages of 25 and 64 who reported working at least 39 weeks per year and at least 30 hours per week. Percentiles were tabulated separately for each LIS file from Waves 4 and 5.³⁵ This provided country- and year-specific percentiles in units of nominal national currency. These were then converted to US dollars using exchange rates from the Penn World Tables.³⁶ These were deflated using the US Consumer Price Index to provide wages in constant 2000 US dollars. The after-tax measures were constructed by multiplying these estimates by one minus the relevant yearly tax rate (see below). Finally, we averaged over years to create the country-specific measures we used in the regressions.

WDI/WIDER estimates

We use GDP data from the World Development Indicators (WDI) database, and Gini coefficients from the World Income Inequality Database, v. 2.0 (WIID), to estimate 20th and 80th income percentiles for all countries.

In the WIID database, the unit of observation is an income inequality study that produced a Gini coefficient. Other variables indicate the country for which the Gini coefficient was computed; the year in which the underlying income survey was fielded; the coverage of the underlying income survey, in terms of the target population, target geography, and target age range; and the definition of the income measure under study. We included only studies of income data collected between 1990 and 2000 which targeted the full population, geography, and age distribution of the country. We also limited attention to studies of income, as opposed to earnings or consumption, since income studies were much more numerous.

We then transformed each Gini coefficient into the standard deviation of log income, assuming that income is lognormal. If income X is lognormally distributed, so $\ln X \sim N(\mu, \sigma^2)$, then the relationship between σ and the Gini coefficient for income, denoted by G , is given by $\sigma = \sqrt{2}\Phi^{-1}\left(\frac{G+1}{2}\right)$, where Φ^{-1} is the inverse of the standard normal cdf (Bendel et al 1989, p. 399). We averaged over all such observations for each country to obtain country-specific standard deviations, denoted by $\bar{\sigma}$.

³⁵ The countries and years for which LIS files were available during waves 4 and 5 were: Australia (1995, 2001), Austria (1994, 1995, 1997, 2000), Canada (1994, 1997, 1998, 2000), Denmark (1995, 2000), France (1994, 2000), Germany (1994, 2000), the Netherlands (1994, 1999), Norway (1995, 2000), Spain (1995, 2000), Sweden (1995, 2000), the UK (1994, 1995, 2000), and the US (1994, 1997, 2000).

³⁶ For the PPP-adjusted wage measures, we applied PPP indexes from the Penn World Tables.

Under log normality, quantiles of income x_α such that $P(X < x_\alpha) = \alpha$ are given by

$$(A1) \quad x_\alpha = \exp(\mu + z_\alpha \sigma),$$

where z_α is the α quantile of a unit normal random variable (Johnson and Kotz, 1970, p. 117). Since $E(X) = \exp(\mu + \sigma^2 / 2)$ under log-normality, we can re-write (A1) as

$$(A2) \quad x_\alpha = E(X) \exp(\sigma z_\alpha - \sigma^2 / 2).$$

To estimate the 20th and 80th percentiles of income using (A2), we used per capita GDP in 1990 and 2000 to estimate $E(X)$ and $\bar{\sigma}$ to estimate σ . The after-tax measures for destination countries which were used in the sorting regressions were constructed by multiplying these estimates by one minus the relevant average tax rate, where averages were taken over the period 1996 to 2000 (see below).

Freeman-Oostendorp data

Freeman and Oostendorp (2000) provide estimates of average annual earnings by occupation and industry for a large cross section of countries, covering the years 1983 to 2003. For most countries, data are available for a just a few years within this span. We restrict our attention to the period 1988 to 1997. For each country, we convert earnings into 2000 US dollars and then take mean earnings across time by occupation and industry. For each country, we take earnings at the 10th percentile to correspond to low-skill wages and earnings at the 80th percentile to correspond to high-skill wages (these percentiles give skill related wage differences that have the highest correlation with the LIS, relative to other percentiles). FO estimate several different versions of their earnings series. We use the series that corresponds to “country-specific calibrations with imputation,” as this was the most highly correlated with the LIS.

Tax data

Tax data are drawn from OECD (various years). These documents provide country-specific annual average tax rates by income level, marital status, and family size for the period 1996 to 2000. We imputed 1996 values to 1994 and 1995 to merge with the annual country-level LIS data. Income levels are defined as a share of the average production worker wage (APW). To 20th percentile wages we applied the tax rate facing workers with income equal to 67 percent of the APW. To 80th percentile wages we applied the tax rate applicable to workers with income equal to 167 percent of the APW. Since the tax treatment of marriage and children varies widely among countries, but we had no data on family status, we applied the tax rate facing single workers with no children.

All wage data used in the analysis are available at:

<http://harrisschool.uchicago.edu/faculty/web-pages/jeffrey-grogger-migration-files.asp>.

Figure 1: Share of emigrants and general population with 13+ years of schooling

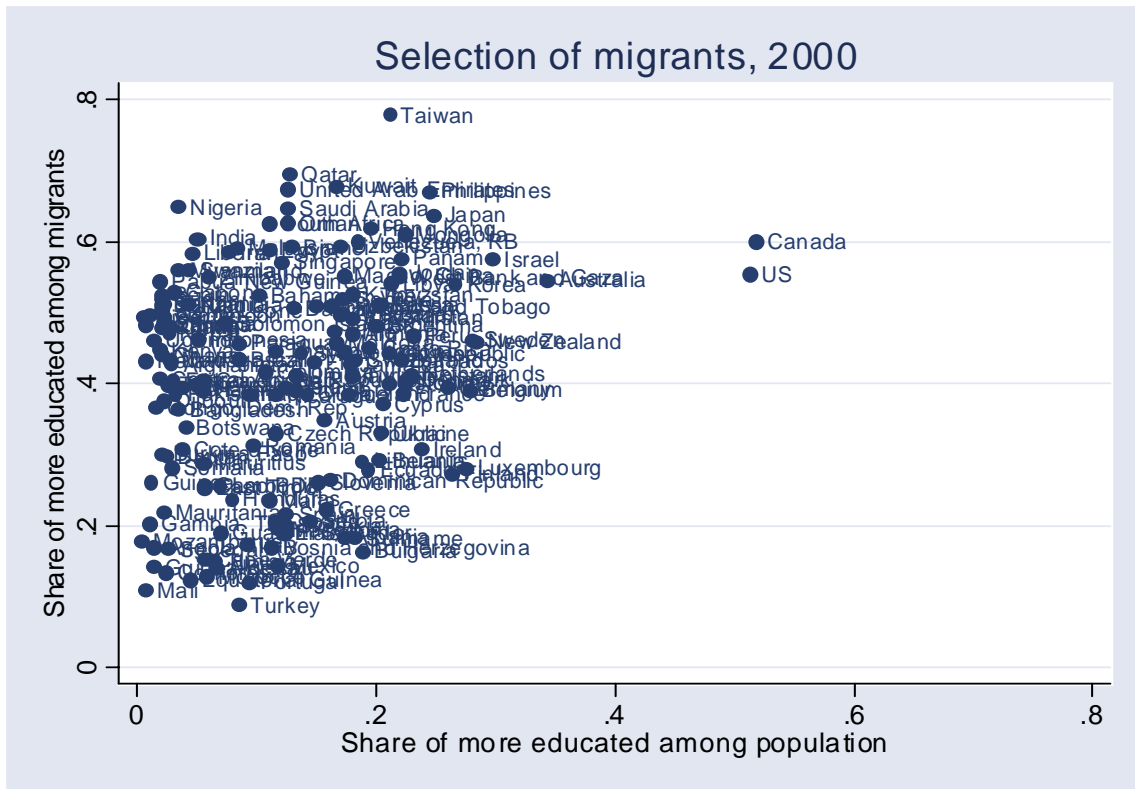


Figure 2: Emigration odds (primary and tertiary educated) by source country, 2000

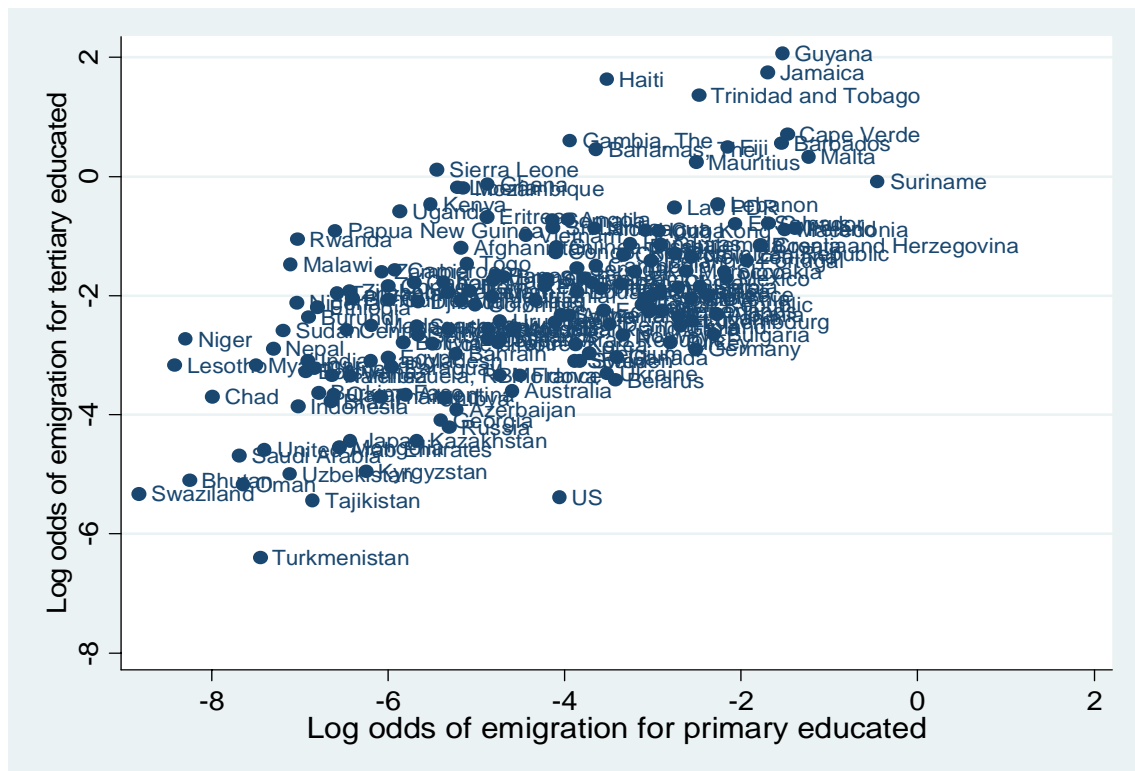


Figure 3a: Log differences in income by country

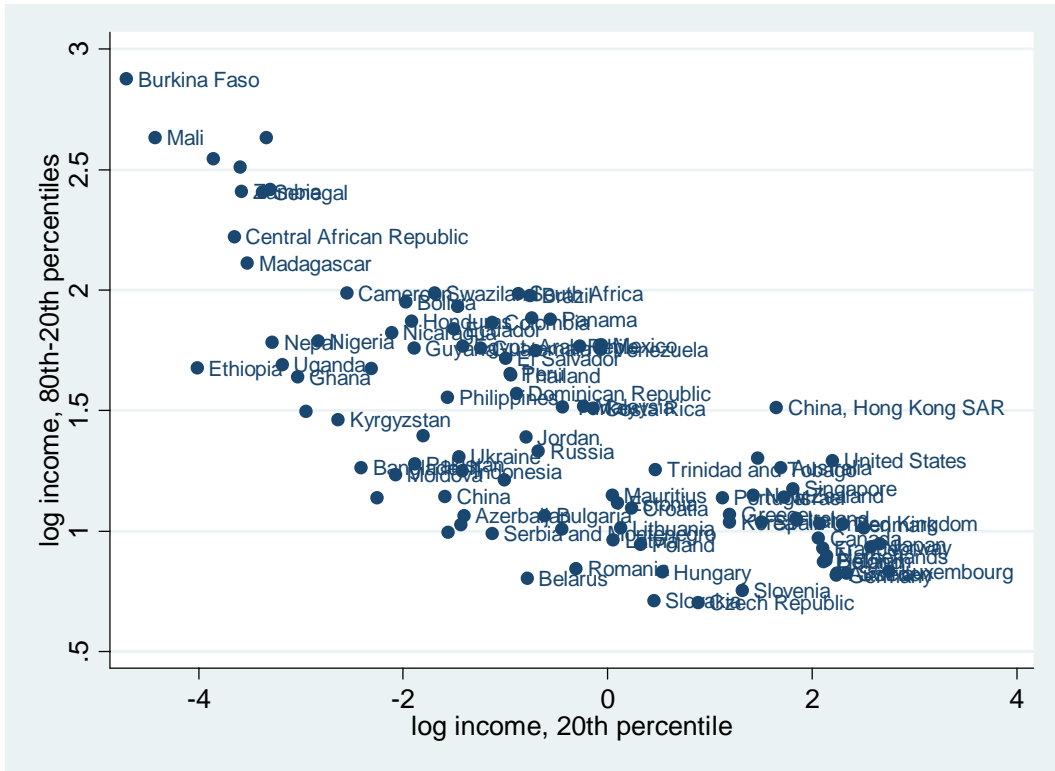


Figure 3b: Absolute differences in income by country (000s USD)

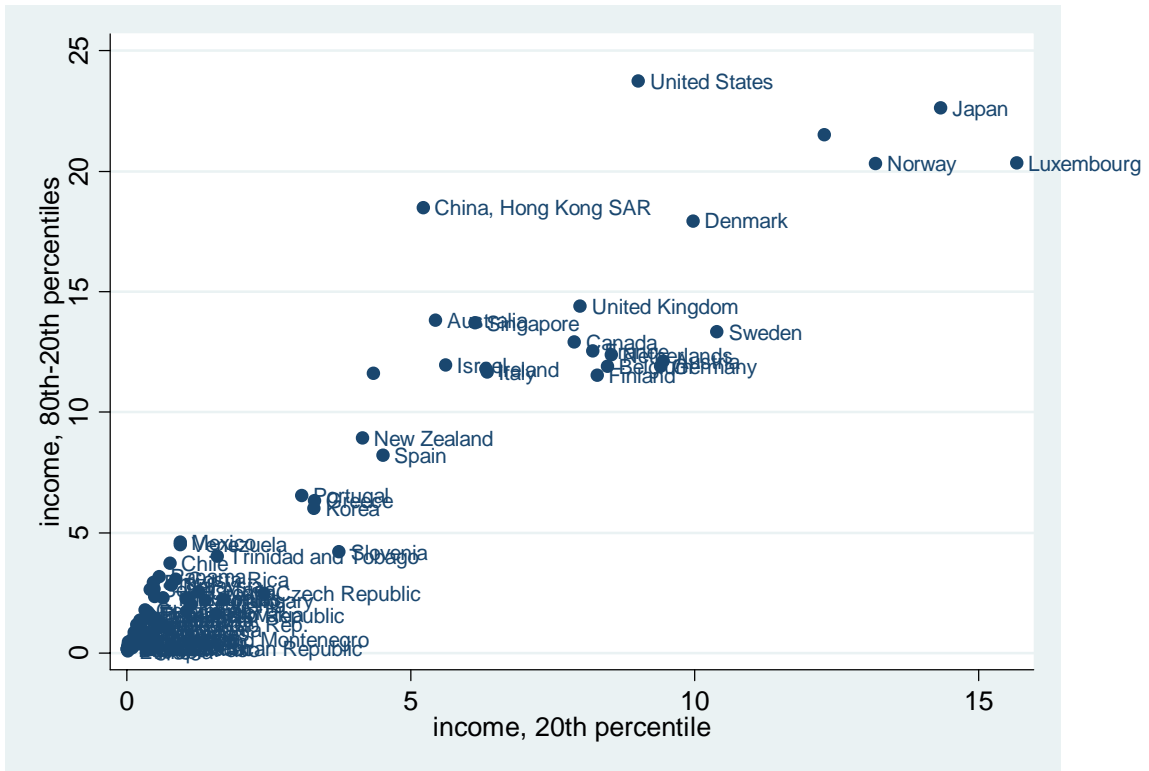


Table 1: Share of OECD immigrants by destination region and education, 2000

Destination Region	Education Group			
	All	Primary	Secondary	Tertiary
North America	0.514	0.352	0.540	0.655
Europe	0.384	0.560	0.349	0.236
Australia & Oceania	0.102	0.088	0.111	0.109
All OECD		0.355	0.292	0.353

Notes: This table shows the share of immigrants in OECD countries by schooling group and destination region for North America (Canada, Mexico the US); Australia & Oceania (Australia, Japan, New Zealand, Korea), and Europe (other OECD members as of 2000).

Table 2: Share of emigrants to OECD by source country and destination region, 2000

Source country	Destination region			
	All OECD	N. America	Europe	Aus. & Oceania
Mexico	0.113	0.219	0.001	0.000
UK	0.053	0.041	0.027	0.206
Italy	0.042	0.027	0.062	0.038
Germany	0.038	0.028	0.049	0.045
Turkey	0.035	0.003	0.085	0.005
India	0.030	0.038	0.023	0.018
China	0.030	0.039	0.009	0.066
Philippines	0.030	0.046	0.007	0.030
Vietnam	0.022	0.032	0.008	0.026
Portugal	0.022	0.011	0.040	0.002
Korea	0.021	0.025	0.002	0.075
Poland	0.020	0.019	0.024	0.010
Morocco	0.019	0.002	0.048	0.000
Cuba	0.015	0.028	0.002	0.000
Canada	0.015	0.025	0.004	0.006

Notes: This table shows the share of immigrants accounted for by the 15 largest source countries for migrants to OECD destination countries.

Table 3: Summary statistics for wage data

A: Destination countries

Pre-tax Wage:	Low-skill	High-skill	Difference	N
Source	(1)	(2)	(3)	(4)
LIS	20.12	41.87	21.76	13
WDI	8.18	21.9	13.71	15
FO	15.71	26.25	10.54	12

Post-tax wage:	Low-skill	High-skill	Difference	N
Source	(1)	(2)	(3)	(4)
LIS	13.02	23.41	10.39	13
WDI	5.22	12.15	6.94	15
FO	10.03	14.56	4.54	12

B: Source countries

Pre-tax wage:	Low-skill	High-skill	Difference	N
Source	(1)	(2)	(3)	(4)
WDI	2.42	6.99	4.57	102
FO	3.97	7.67	3.71	101

Notes: See appendix for details on construction of wage variables.

Table 4: Regression results from linear-utility model

Equation:	Scale	Selection	Sorting	Sorting	Sorting	Sorting
Wage data source:	WDI	WDI	WDI	WDI	LIS	LIS
Variable	(1)	(2)	(3)	(4)	(5)	(6)
$W_h^j - W_s^j$	0.018 (0.029)					
$(W_h^3 - W_h^1) - (W_s^3 - W_s^1)$		0.072 (0.013)				
$(W_h^3 - W_h^1)$, pre-tax			0.060 (0.026)		0.026 (0.013)	
$(W_h^3 - W_h^1)$, post-tax				0.103 (0.045)		0.048 (0.022)
Anglophone dest.	1.451 (0.873)	0.567 (0.183)	0.838 (0.183)	0.636 (0.256)	0.817 (0.193)	0.678 (0.241)
Common language	0.648 (0.293)	1.268 (0.248)	0.355 (0.137)	0.352 (0.139)	0.331 (0.125)	0.332 (0.124)
Contiguous	0.880 (0.401)	-0.384 (0.373)	-1.005 (0.229)	-1.007 (0.237)	-1.108 (0.230)	-1.097 (0.240)
Longitude diff.	0.003 (0.004)	-0.009 (0.003)	0.004 (0.002)	0.004 (0.002)	0.005 (0.003)	0.005 (0.003)
Log distance	-1.152 (0.171)	0.676 (0.131)	-0.245 (0.092)	-0.259 (0.097)	-0.273 (0.107)	-0.279 (0.111)
LT colonial rel.	2.159 (0.411)	-0.711 (0.193)	-0.391 (0.176)	-0.445 (0.161)	-0.505 (0.150)	-0.550 (0.137)
ST colonial rel.	2.641 (0.601)	-0.395 (0.431)	-0.129 (0.256)	-0.187 (0.257)	-0.195 (0.276)	-0.224 (0.276)
Visa waiver	0.589 (0.314)	-0.299 (0.135)	0.335 (0.164)	0.364 (0.172)	0.440 (0.200)	0.471 (0.203)
Schengen sig.	0.058 (0.337)	0.402 (0.166)	0.430 (0.250)	0.403 (0.252)	0.528 (0.295)	0.507 (0.304)
Asylee share	-1.221 (3.698)	-2.512 (0.818)	-3.590 (0.901)	-3.635 (0.709)	-3.998 (0.929)	-4.007 (0.810)
Observations	2786	1393	1393	1393	1214	1214
R-squared	0.44	0.47	0.61	0.61	0.63	0.63
Clusters	15	15	15	15	13	13

Notes: Robust standard errors in parentheses. In addition to the variables shown, all regressions include a dummy variable equal to one if the destination-country statistical office explicitly codes a primary education category. The scale regression also includes a dummy equal to one for the tertiary skill-group observations and interactions between that dummy and all variables shown. The sorting regressions include a full set of source-country dummies.

Table 5: Regression results from log-utility model

Equation:	Scale	Selection	Sorting	Sorting	Sorting	Sorting
Wage data source:	WDI	WDI	WDI	WDI	LIS	LIS
Variable	(1)	(2)	(3)	(4)	(5)	(6)
$\ln W_h^j - \ln W_s^j$	-0.435 (0.087)					
$(\delta_h^3 - \delta_s^1)$		-1.307 (0.186)				
δ_h^3 , pre-tax			3.929 (0.767)		5.338 (0.886)	
δ_h^3 , post-tax				3.342 (0.761)		4.146 (1.297)
Anglophone dest.	1.466 (0.857)					
Common language	1.315 (0.213)					
Contiguous	0.656 (0.301)					
Longitude diff.	-0.008 (0.003)					
Log distance	-0.530 (0.180)					
LT colonial rel.	1.912 (0.485)					
ST colonial rel.	2.185 (0.413)					
Visa waiver	-0.793 (0.259)					
Schengen sig.	-0.523 (0.350)					
Asylee share	-2.065 (3.257)					
Observations	2786	1393	1393	1393	1214	1214
R-squared	0.29	0.17	0.40	0.38	0.43	0.38
Clusters	15	15	15	15	13	13

Notes: Robust standard errors in parentheses. In addition to the variables shown, all regressions include a dummy variable equal to one if the destination-country statistical office explicitly codes a primary education category. The sorting also regressions include a full set of source-country dummies.

Table 6: Key wage coefficients based on alternative wage measures

A. WDI wages, PPP-adjusted		
Equation: Variable	Selection (1)	Sorting (2)
$(W_h^3 - W_h^1) - (W_s^3 - W_s^1)$	0.108 (0.016)	
$(W_h^3 - W_h^1)$, post-tax		0.082 (0.054)
Observations	1379	1379
R-squared	0.50	0.60
Clusters	15	15
B. LIS wages, PPP adjusted		
Equation: Variable	Selection (1)	Sorting (2)
$(W_h^3 - W_h^1)$, post-tax		0.047 (0.024)
Observations		1202
R-squared		0.63
Clusters		13
C. Freeman-Oostendorp wages		
Equation: Variable	Selection (1)	Sorting (2)
$(W_h^3 - W_h^1) - (W_s^3 - W_s^1)$	0.082 (0.016)	
$(W_h^3 - W_h^1)$, post-tax		0.072 (0.045)
Observations	1093	1093
R-squared	0.49	0.63
Clusters	12	12
D. Freeman-Oostendorp wages, PPP-adjusted		
Equation: Variable	Selection (1)	Sorting (2)
$(W_h^3 - W_h^1) - (W_s^3 - W_s^1)$	0.064 (0.011)	
$(W_h^3 - W_h^1)$, post-tax		0.091 (0.039)
Observations	1059	1059
R-squared	0.49	0.64
Clusters	12	12

Notes: Robust standard errors in parentheses. In addition to variables shown, all regressions include all variable shown or discussed in the note to Table 4.

Table 7: Additional Selection and Sorting Regressions

Equation:	Selection	Sorting	Sorting	Selection	Sorting	Sorting	Selection	Sorting	Sorting	Sorting	Sorting
Wage data source:	WDI	WDI	LIS	WDI	WDI	LIS	WDI	WDI	LIS	WDI	LIS
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
$(W_h^3 - W_h^1) - (W_s^3 - W_s^1)$	0.065 (0.013)			0.078 (0.012)			0.082 (0.012)				
$(W_h^3 - W_h^1)$, post-tax		0.121 (0.054)	0.052 (0.025)		0.127 (0.053)	0.058 (0.015)		0.121 (0.047)	0.068 (0.019)	0.106 (0.039)	0.050 (0.018)
Relative university quality				0.000 (0.001)	-0.002 (0.002)	-0.004 (0.002)					
Log emigrant stock 1990							-0.148 (0.031)	-0.039 (0.055)	-0.157 (0.058)		
Observations	1393	1393	1214	1348	1348	1169	963	963	823	2338	2044
R-squared	0.40	0.57	0.57	0.49	0.63	0.65	0.59	0.64	0.69	0.59	0.62
Clusters	15	15	13	14	14	12	15	15	13	15	13

Notes: Robust standard errors in parentheses. In addition to variables shown, all regressions include all variables shown or discussed in the note to Table 4. Columns (1)-(3) report regressions in which we redefine skilled migrants to be migrants with either secondary or tertiary education; columns (4)-(9) differ from the corresponding regressions in Table 4 only by in the inclusion of the indicated regressor; and columns (10) and (11) re-estimate the sorting regression using the full sample of source-destination pairs for which destination wage data are available.

Table 8: Wage coefficients from selection and sorting regressions from samples that omit one destination

<i>A. Selection regressions; WDI wage measures</i>															
Omitted destination	AUS	AUT	CAN	DEU	DNK	ESP	FIN	FRA	GBR	IRL	NLD	NOR	NZL	SWE	USA
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(15)	(16)
$(W_h^3 - W_h^1) - (W_s^3 - W_s^1)$	0.066 (0.013)	0.071 (0.013)	0.068 (0.012)	0.071 (0.013)	0.078 (0.012)	0.078 (0.012)	0.072 (0.013)	0.075 (0.013)	0.072 (0.013)	0.079 (0.012)	0.074 (0.013)	0.066 (0.013)	0.067 (0.013)	0.074 (0.013)	0.073 (0.015)
Observations	1296	1294	1294	1303	1299	1298	1306	1293	1293	1348	1293	1295	1301	1292	1297
R-squared	0.47	0.47	0.46	0.47	0.47	0.48	0.46	0.48	0.47	0.49	0.47	0.48	0.48	0.48	0.44
Clusters	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14
<i>B. Sorting regressions; WDI post-tax wage measures</i>															
Omitted destination	AUS	AUT	CAN	DEU	DNK	ESP	FIN	FRA	GBR	IRL	NLD	NOR	NZL	SWE	USA
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(15)	(16)
$(W_h^3 - W_h^1)$, post-tax	0.094 (0.043)	0.110 (0.040)	0.112 (0.035)	0.111 (0.038)	0.122 (0.036)	0.112 (0.039)	0.113 (0.039)	0.103 (0.040)	0.108 (0.039)	0.141 (0.037)	0.112 (0.039)	0.080 (0.040)	0.095 (0.040)	0.134 (0.039)	0.122 (0.054)
Observations	2166	2175	2170	2206	2180	2176	2204	2161	2152	2287	2156	2177	2178	2160	2184
R-squared	0.40	0.39	0.39	0.39	0.39	0.41	0.39	0.40	0.40	0.41	0.39	0.41	0.42	0.41	0.38
Clusters	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14
<i>C. Sorting regressions; LIS post-tax wage measures</i>															
Omitted destination	AUS	AUT	CAN	DEU	DNK	ESP	FRA	GBR	IRL	NLD	NOR	SWE	USA		
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)		
$(W_h^3 - W_h^1)$, post-tax	0.031 (0.020)	0.051 (0.019)	0.053 (0.016)	0.051 (0.017)	0.055 (0.018)	0.051 (0.018)	0.047 (0.018)	0.052 (0.018)	0.067 (0.015)	0.051 (0.018)	0.037 (0.014)	0.055 (0.018)	0.071 (0.033)		
Observations	1872	1881	1876	1912	1886	1882	1867	1858	1993	1862	1883	1866	1890		
R-squared	0.43	0.41	0.41	0.41	0.41	0.43	0.42	0.42	0.43	0.41	0.43	0.43	0.40		
Clusters	12	12	12	12	12	12	12	12	12	12	12	12	12		

Notes: Robust standard errors in parentheses. In addition to variables shown, all regressions include all variable shown or discussed in the note to Table 4.

Table 9: Fixed migration costs and migrant stocks for selected source and destination countries

Destination: Source	Australia	Canada	France	Germany	UK	US
Canada	59.94 21,375		53.84 8,910	48.66 7,998	33.7 49,954	10.6 715,825
China	80.28 117,170	56.8 287,820	90.83 24,547	88.57 26,069	87.73 33,380	55.6 841,699
Colombia	89.88 3,083	69.4 11,725	72.84 6,136	71.72 5,895	71.43 8,928	34.67 402,935
Dominican Republic	133.45 38	62.58 3,225	90.62 238	59.85 3,188	95.37 350	6.62 527,520
El Salvador	50.52 6,314	26.82 27,780	83.06 328	82.21 332	89.85 302	-0.16 619,185
Germany	34.69 102,219	19.94 163,880	14.8 109,425		13.76 164,165	21.83 646,815
Guatemala	105.82 157	45.01 8,880	83.78 346	80.14 448	96.95 212	8.86 341,590
Ireland	19.66 45,365	22.87 24,520	34.63 3,845	13.76 13,284	-28.28 420,102	16.55 148,680
Italy	27.4 216,316	14.95 312,185	13.37 371,714	10.91 456,000	31.34 86,876	23.3 461,085
Jamaica	76.79 680	-1.73 103,265	67.18 299	51.03 872	-9.31 124,313	-2.98 449,795
Korea, Rep.	66.14 25,160	48.94 50,860	73.86 6,164	60.2 12,226	72.08 7,434	29.06 676,640
Mexico	117.03 870	59.6 24,795	90.46 3,064	82.5 4,029	90.62 3,558	0 6,374,825
Philippines	65.41 78,105	46.64 191,615	83.42 4,767	68.31 12,539	67.99 34,782	29.58 1,163,555
Poland	45.63 52,887	26.38 154,525	27.95 91,122	5.21 198,000	42.89 33,661	29.83 399,165
Portugal	46.13 13,329	9.32 143,145	-2.55 538,106	8.81 113,216	31.08 26,006	19.23 187,645
Spain	66.39 11,972	63.67 9,695	15.77 308,500	27.11 109,613	42.88 40,592	50.85 73,835
Turkey	63.17 26,160	68.03 13,045	40.1 133,890	7.95 1,272,000	53.27 36,754	60.36 64,780
United Kingdom	12.39 966,139	11.87 580,250	23.02 61,317	17.29 90,000		36.99 613,930
Vietnam	39.56 128,666	31.02 127,590	38.45 58,570	45.46 43,105	61.75 19,137	21.01 807,305

Note: Top figure in each cell is fixed migration cost in 000s of annual 2000 USD; bottom figure is emigrant stock.

Table 10: Decomposition of the immigrant skills gap

Destination	Share of immigrant skills gap explained by:													Share explained (14)
	Mean immigrant skills (1)	Immigrant skills gap (2)	Wage difference (3)	English dest. (4)	Common off. lang. (5)	Contiguous (6)	Long. diff. (7)	Log distance (8)	Colony, LT (9)	Colony, ST (10)	Visa waiver (11)	Schengen (12)	Share asylees (13)	
Australia	1.29	0.72	0.84	0.00	-0.01	-0.03	-0.16	0.22	-0.01	0.00	-0.01	0.00	-0.07	0.78
Austria	-0.54	2.55	0.36	0.25	0.03	0.02	0.07	-0.08	0.00	0.00	-0.04	-0.02	0.39	0.98
Canada	1.30	0.71	0.99	0.00	-0.06	-0.02	0.00	0.02	-0.01	0.00	-0.04	0.00	-0.13	0.76
Denmark	-0.51	2.52	0.28	0.25	0.04	0.00	0.07	-0.07	0.00	0.00	-0.04	-0.02	0.36	0.87
Finland	-0.75	2.76	0.38	0.23	0.03	0.00	0.07	-0.06	0.00	0.00	-0.04	-0.02	0.09	0.68
France	-0.10	2.11	0.44	0.30	0.02	0.02	0.08	-0.08	0.02	0.00	-0.05	-0.02	0.22	0.96
Germany	-0.31	2.32	0.44	0.27	0.03	0.03	0.08	-0.09	0.00	0.00	-0.05	-0.03	0.05	0.76
Ireland	1.39	0.62	1.52	0.00	0.00	0.00	0.27	-0.37	-0.01	0.00	-0.31	0.00	0.83	1.95
Netherlands	-0.54	2.55	0.32	0.25	0.03	0.00	0.07	-0.07	0.00	0.00	-0.04	-0.02	0.55	1.10
N. Zealand	0.71	1.30	0.64	0.00	0.00	-0.02	-0.13	0.14	0.00	0.00	-0.04	0.00	-0.26	0.32
Norway	1.15	0.86	0.53	0.74	0.10	0.01	0.21	-0.17	-0.01	0.00	-0.12	-0.06	0.40	1.64
Spain	-0.03	2.04	0.50	0.31	0.02	0.00	0.08	-0.06	0.03	0.00	-0.05	-0.03	-0.06	0.74
Sweden	0.28	1.73	0.55	0.37	0.05	0.00	0.11	-0.09	0.00	0.00	-0.06	-0.03	0.36	1.27
UK	0.40	1.61	0.35	0.00	0.00	-0.01	0.10	-0.09	0.07	0.00	-0.06	0.00	0.22	0.58
US	2.01	0.00												
Mean	0.38	1.63	0.58	0.21	0.02	0.00	0.07	-0.06	0.01	0.00	-0.07	-0.02	0.21	0.96

Note: Results are based on the model reported in column (4) of Table 4.