Income Maximization and the Selection and Sorting of International Migrants

Jeffrey Grogger, University of Chicago and NBER

Gordon H. Hanson, UCSD and NBER

February 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 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 level 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, Harvard, LSE, Princeton, UCSD, UC Irvine, UCL, and the 2007 NBER Summer Institute. Any errors are ours alone.

1. Introduction

International migration is a potentially important mechanism for global economic integration. As of 2005, individuals residing outside their country of birth accounted for 3% of the world's population. Most of those migrants left home bound for rich nations. The UN estimates that in 2005, 40.9% of the global emigrant population resided in just eight rich economies,[1] with 20.2% living in the U.S. alone. In major destination countries, the number of foreign born is rising, reaching 12.5% of the total population in the U.S., 11.2% in Germany, 10.5% in France, and 8.2% in the U.K.

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) on emigration by schooling group, 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; that is, that they are more educated than their non-migrant counterparts. This observation has renewed interest in the impact of brain drain on developing economies.[2]

A second – and perhaps less appreciated – feature of international migration is the sorting of emigrants across destinations. Countries with high rewards to skill 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 U.S. and Canada, where skill-related wage differences are relatively large, receives 51 percent of the OECD's immigrants, but

[1] These countries are the US, Germany, France, Canada, the UK, Spain, Australia, and Italy. Freeman (2006) notes that Russia and several Middle Eastern countries also receive large numbers of immigrants.
[2] Recent empirical work on brain drain includes Adams (2003), Beine, Docquier and Rapoport (2001, 2007, 2008), Docquier and Rapoport (2007), and Kapur and McHale (2005).

66 percent of its immigrants with tertiary schooling. Europe, where skill-related wage differences are relatively small, receives 38 percent of the OECD's immigrants, but only 24 percent of its tertiary-schooled immigrants. Europe's failure to receive educated migrants may explain its recent efforts to attract skilled foreigners.[3]

In this paper, we develop and estimate a simple model of migration based on the Roy (1951) income maximization framework. The Roy model, which is the foundation for a large body of migration research (Borjas, 1999), implies that the selectivity of migrants and their sorting across destinations should depend on cross-country differences in the reward to skill. Our version of the model predicts that an increase in the reward to skill in a destination should cause immigration from source countries to rise and the mix of migrants to become more skilled.

The model delivers estimating equations for the scale of migration, the selection of migrants in terms of schooling, and the sorting of migrants across destinations by schooling. While the three equations estimate a common coefficient on earnings, they differ in terms of the data they require and the assumptions one must impose regarding migration costs. The scale regression requires data on earnings by schooling level in the source and destination and an assumption that the determinants of fixed migration costs are observable. The selection regression differences out fixed migration costs. The sorting regression does so as well, and also controls for source-specific determinants of migration, including source-country earnings. We analyze newly available data from Beine, Docquier and Rapoport (2007) on the stock of migrants by education level from 192 source countries residing in OECD destination countries as of 2000.

---

[3] See "Not the Ace in the Pack: Why Europe Loses in the Global Competition for Talent," *The Economist*, October 25, 2007.

To preview the findings, the data strongly support income maximization. In the scale regression, migration is increasing in the level earnings difference between the destination and the source, although the estimated effect of earnings appears to be attenuated due to omitted fixed costs of migration. In the selection and sorting regressions, which difference out fixed costs, the relative stock of more-educated migrants is larger in destinations with greater skill-related earnings differences. We also find post-tax earnings are a stronger correlate of migration than pre-tax earnings, consistent with migrants weighing tax treatment. Further results address the role of language, distance, migration policy, historical relationships, and lagged migration.

One 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 country with high returns to skill to a destination with low returns to skill should be negatively selected. Although the Borjas (1987) framework performs well in explaining migration from Puerto Rico to the U.S. (Ramos, 1992; Borjas, 2006), it does less well elsewhere. Migrants from Mexico to the U.S. are drawn from the middle of the skill distribution, even though returns to skill are higher in Mexico than the US.[4] Figure 1 shows that OECD-bound migrants are positively selected, even though many are from countries where returns to skill exceed those in the OECD.

Our results suggest that one explanation for positive selection is that migrants are influenced by skill-related differences in wage levels, rather than relative returns to skill, which is consistent with cross-country differences in labor productivity being a dominant factor in why labor moves across borders. In a world where wage level differences

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

matter, high-skill workers from low-wage countries may have a strong incentive to migrate, even if returns to skill are high in the source country. We also estimate an alternative version of the income maximization model in which relative returns, rather than wage level differences, influence migrant selectivity.[5] The data reject this model.

Our results on scale and selection are consistent with Rosenzweig (2007), who examines legal migration to the U.S. and finds that source-country emigration rates are decreasing in source-country labor productivity. This is comparable to our finding that migration is increasing in the destination-source earnings difference by skill group.[6] Relative to his work, we extend the analysis to multiple destinations, which enables us to analyze sorting as well as scale and selectivity and to account for the relative contribution of earnings and migration costs to international migration. We use our scale regression to estimate the fixed costs of migration between 102 source countries and 15 destination countries, finding that these costs are large, often an order of magnitude greater than source-country earnings for low-skilled workers. We use our selection regression to decompose emigrant selectivity into components attributable to wages differences and components attributable to migration costs by source region and income level.

A second contribution of the paper is to establish the independence of migrant selection and migrant sorting. While the selectivity of migration by skill depends on the reward to skill in the source country, among other factors, the sorting of migrants by skill does not. Positive sorting is a general implication of income maximization. We provide the first evidence on the sorting of international migrants across destinations; previous

---

5 Other work on bilateral migration tends to use log per capita GDP to measure wages often with controls for income inequality. See Volger and Rotte (2000), Pedersen, Pytlikova, and Smith (2004), Hatton and Williamson (2005), Mayda (2005), and Clark, Hatton, and Williamson (2007).

6 In related work, Rosenzweig (2006) finds that the number of students who come to the U.S. for higher education and who then stay in the U.S. are decreasing in labor productivity in the source country.

studies of sorting focused on internal US migration (Borjas, Bronars, and Trejo 1992; Dahl 2002). We use our 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 U.S. and Canada receive more skilled immigrants than other OECD destinations.

In section 2, we present a simple model of international migration and derive the estimating equations. In section 3, we describe our data. In section 4, we give the estimation results. Section 5 offers concluding remarks.

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 be[7]

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

where $\exp(\mu_h)$ is the wage paid to workers with primary education, δ_h^2 is the return to secondary education, δ_h^3 is the return to tertiary education, and D_{is}^j is a dummy variable indicating whether individual 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 common to all individuals

[7] 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 skill group and source country, it is not possible to address within group heterogeneity in skill.

who move from s to h, f_{sh}; and a component that varies by skill group, g^{j}_{sh} (which may be positive or negative), such that

(2) $$C^{j}_{ish} = f_{sh} + g^{1}_{sh}D^{1}_{i} + g^{2}_{sh}D^{2}_{i} + g^{3}_{sh}D^{3}_{i}.$$

Migration costs are influenced by the linguistic and geographic distance between the source and the destination and by destination-country immigration policies. The impacts 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^{j}_{ish} such that

(3) $$U^{j}_{ish} = \alpha(W^{j}_{ih} - C^{j}_{ish}) + \varepsilon^{j}_{ish},$$

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^{j}_{ish} follows an i.i.d. extreme value distribution, we can write the log odds of migrating to destination country h versus staying in the source country s for members of skill group j as[8]

[8] 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 more complicated nesting structures, in which we examine only the OECD branch of the decision tree (one such structure would be in which individuals first choose to migrate or not migrate, migrants then choose either OECD or non-OECD sets of destination countries, and sub-migrants then choose among destinations within these sets). Alternatively, one might imagine that

$$\text{(4)} \qquad \ln \frac{E_{sh}^{j}}{E_{s}^{j}} = \alpha(W_{h}^{j} - W_{s}^{j}) - \alpha f_{sh} - \alpha g_{sh}^{j}$$

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 $W_{h}^{j} = e^{\mu_h + \delta_h^j}$ (McFadden 1974). Equation (4) speaks to the scale of migration. It says that income maximization, together with our 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:

$$\text{(5)} \qquad \ln \frac{E_{sh}^{3}}{E_{sh}^{1}} - \ln \frac{E_{s}^{3}}{E_{s}^{1}} = \alpha[(W_{h}^{3} - W_{s}^{3} - g_{sh}^{3}) - (W_{h}^{1} - W_{s}^{1} - g_{sh}^{1})].$$

The first term on the left side of (5) is a measure of the skill distribution 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, and 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 distributions between emigrants (from s to destination 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 countries,

---

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).

net of skill-varying migration costs, is greater for high-skill workers. Emigrants should be negatively selected if the net source-destination wage difference is greater for low-skill workers. Note that fixed costs f_{sh} do not appear in the selection equation (5). Differencing between skill groups has eliminated them from the expression.

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

$$\text{(6)} \qquad \ln\frac{E_{sh}^{3}}{E_{sh}^{1}} = \alpha(W_{h}^{3} - W_{h}^{1}) - \alpha(g_{sh}^{3} - g_{sh}^{1}) + \tau_{s}$$

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) expresses the key implication of utility maximization in the presence of multiple destinations. Since α > 0, 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 *level* 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). It is useful to compare these two models theoretically and empirically.

To do so, consider a log utility model where wages and migration costs are as before, but utility is given by

$$(7) \qquad U^j_{ish} = (W^j_{ih} - C^j_{ish})^\lambda \exp(\upsilon^j_{ish})$$

where $\lambda > 0$ and υ^j_{ish} 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

$$(8) \qquad \ln \frac{E^j_{sh}}{E^j_s} = \lambda(\ln W^j_h - \ln W^j_s) - \lambda m^j_{sh}$$

$$(9) \qquad \ln \frac{E^3_{sh}}{E^1_{sh}} - \ln \frac{E^3_s}{E^1_s} = \lambda(\delta^3_h - \delta^3_s) - \lambda(m^3_{sh} - m^1_{sh})$$

$$(10) \qquad \ln \frac{E^3_{sh}}{E^1_{sh}} = \lambda\delta^3_h - \lambda(m^3_{sh} - m^1_{sh}) + \rho_s$$

where $m^j_{sh} = \left(f_{sh} - g^j_{sh}\right) / W^j_h$ and $\rho_s = \ln(E^3_s / E^1_s) - \lambda\delta^3_s$.[9] 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)).

In the log utility model, differencing between skill groups does not in general eliminate either fixed or skill-varying migration costs from the selection or sorting equations in (9) and (10). In the special case where skill-varying costs are proportional to wages, such that $g^j_{sh} = \pi_{sh} W^j_h$, differencing between skill groups eliminates skill-varying

---

[9] In deriving (8), we use the approximation that ln(W-C) ≈ lnW – C/W for sufficiently small C/W. Equation (9) follows from the fact that $\ln W^3_h - \ln W^1_h = \delta^3_h$.

costs, but not fixed costs. Since much of the literature has focused on models where skill-varying costs are assumed to be proportional to wages and fixed costs are assumed to be zero, it represents 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 the definition of W_h^j into the right side of (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

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

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

The two models make similar predictions about migrant selectivity in the context of typical north-to-north migration, where 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 $W_h^1 >> W_s^1$. 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).[10]

The evidence suggests that returns to schooling tend to be higher in developing countries than in the U.S. or Europe (Psacharopoulos and Patrinos, 2004; Hanushek and Zhang, 2006), so the log utility/proportional cost model implies that emigrants from developing countries should tend to be negatively selected. This prediction is clearly at odds with Figure 1. However, the linear utility model could be consistent with Figure 1, so along as productivity differences across countries dominate differences in the returns to schooling (or skill-specific migration costs are higher for low-skill workers).

While many studies have tested for the selectivity of migrants, fewer analysts have examined migrant sorting across multiple destinations. Borjas, Bronars, and Trejo (1992) develop a theoretical model that predicts sorting on the basis of destination returns to skill. They and Dahl (2002) estimate empirical models of sorting using data on internal migration in the U.S. There have been no studies of migrant sorting in the context of international labor flows.

One point that seems to have escaped the theoretical literature is that selection and sorting are logically independent. In terms of our model, sorting between destinations h and k depends on the sign of

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

[10] 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^3_{sh}, 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.

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 $\Delta^{hk} > 0$, then destination h should receive more highly skilled migrants than destination k. This should hold whether emigrants from s to both h and k are positively selected ($\Delta^h > 0, \Delta^k > 0$), negatively selected ($\Delta^h < 0, \Delta^k < 0$), or even bimodally selected ($\Delta^h < 0$ and $\Delta^k > 0$ or vice-versa).

C. Estimation

Although the discussion until now has been cast in terms of population magnitudes, it is straightforward to derive an estimating equation 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 $g_{sh}^j = x_{sh}\theta^j$.[11] The empirical version of the scale equation is

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

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

[11] The analysis is partial equilibrium in nature and cannot be used to examine how bilateral migration flows affect the wage structure in destination countries.

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

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

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 to be tested in each regression is that $\alpha > 0$, as utility maximization requires. Indeed, if the models are properly specified, all three equations should yield similar estimates of α. However, an important difference among these specifications is the treatment of fixed costs. 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. In contrast, 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 potential source countries. The sorting equation requires only destination-country wage data, increasing the number of source countries that can be used to estimate the model. Additionally, measurement error may be lower in the destination countries, comprised of OECD members, than in source countries, which include the developing world.

Finally, we estimate the log-utility model so as to provide a direct comparison with the linear-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

$$(15)\qquad \ln \frac{\hat{E}^j_{sh}}{\hat{E}^j_s} = \lambda(\ln W^j_h - \ln W^j_s) + x_{sh}\theta + \eta^j_{sh},$$

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

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

where we have assumed that $-\lambda\pi_{sh} = x_{sh}\theta$. As above, a test for income maximization amounts to a test for $\lambda > 0$, and if the models are properly specified, all three equations should yield similar estimates of λ.

3. Data and Empirical Setting

In the introduction we presented data on skill-specific migration rates which showed evidence of positive selection. They also showed evidence of sorting across multiple destinations of the type predicted by income maximization. Those data are from Docquier and Marfouk (2006). We base our regression analysis on an updated version of these data from Beine, Docquier, and Rapoport (2007; hereafter, BDR).

BDR tabulate data on stocks of emigrants by source and destination country. In collaboration with the national statistical offices of 20 OECD countries, they estimate the population in each OECD country of immigrants 25 years and older by source country and education level. In some of the 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 in a comparable manner at a finer level of detail.

A. Measurement of Emigrant Stocks

Aggregating data from multiple destination countries raises several comparability issues. The first involves the definition of immigrants. 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.

Measuring education levels poses several problems. In Belgium and Italy, the statistical office reports aggregate immigrant counts but does not disaggregate by education level. BDR impute the skill distribution of immigrants in such cases using data from household labor-force surveys, but 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. 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.

Finally, 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 2000 in the United States, the share of foreign-born individuals 25-64 years old with tertiary education who stated they were not in school was 86.4%. 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. Finally, Figure 2 plots the log odds of emigration for the tertiary educated against the log odds of emigration for the primary educated. Nearly all points lie above the 45-degree line, indicating that the log odds of emigration is higher for the more educated, as is consistent with emigrants being positively selected in terms of schooling.

B. Wage Measures

The key explanatory variables in our regression models are functions of skill-group-specific wages in the source and destination countries. Ideally, we would estimate wages by broad education category from the same sources used by DM. 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.

[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.

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 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.[14],

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 data for a large sample of source countries.[15] We rely on two sources of aggregate data to construct the source-destination wage difference measures needed to estimate the scale and sorting regressions.

---

[13] In a previous version of this paper (Grogger and Hanson, 2007), we experimented with alternative measures of wage differences based on various measures of low-skill wages and different measures of the return to skill (the standard deviation of income, the ratio of income in the 80th to 20th percentiles, and the Gini coefficient). All alternatives we considered generated results similar to those we report in this paper.
[14] The years corresponding to each country are as follows: Australia (1995, 2001), Austria (1994, 1995, 1997, 2000), Canada (1994, 1997, 1998, 2000), Denmark (1995, 2000), France (1994, 2000), Germany (1004, 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).
[15] The IPUMS-International study provides samples of Census data for 26 countries, but many important sources and destinations for migrants are not included.

One source combines Gini coefficients from the WIDER World Income Inequality Database with per capita GDP from the World Development Indicators (hereafter, WDI). Under the assumption that income has a log normal distribution, Gini coefficients can be used to estimate the variance of log income.[16] Using per capita GDP to measure mean wages, we can then construct estimates of the 20th and 80th percentiles of wages (see note 17), which we are able to do for 102 source countries and 15 destination countries.[17]

A second source uses data from Freeman and Oostendorp (2000; hereafter FO), who have collected information on earnings by occupation and industry from the 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. We choose these deciles because they 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.

---

[16] Suppose log income is normally distributed with mean μ and variance σ. Given an estimate of the Gini coefficient, G, the standard deviation of log income is given by $\sigma = \sqrt{2}\Phi^{-1}\left(\frac{G+1}{2}\right)$. Note further that the value of log income at the α quantile is given by $\mu \exp(\sigma z_\alpha - \sigma^2/2)$, where z_α is the α quantile of N(0,1).

[17] We restricted attention to Gini coefficients computed from income data over the period 1990-2000, where the underlying sample was drawn from the country's full population. For each country, we averaged over all Gini coefficients that satisfied those criteria. GDP per capita is averaged over the period 1990 to 2000 and expressed in constant 2000 dollars.

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.

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). To 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.[18] In both cases, the tax rate includes income taxes net of benefits plus both sides of the payroll tax. After-tax wage differences are only about half as large as pre-tax differences.

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.

---

[18] 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.

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.[19] To avoid confounding destination-country skilled-unskilled wage differences 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 a source and destination country. Relatedly, proximity may make illegal immigration less costly, thereby increasing the relative migration of less-educated individuals. 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), benefiting lower-income individuals disproportionately (Orrenius and Zavodny, 2005; McKenzie and Rapoport, 2006). Networks may be stronger between countries that share a common colonial heritage, for which we control using CEPII's indicators of whether a pair of countries have short or long colonial histories. We also control for

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

migrant networks using lagged migration, measured as the total stock of emigrants from a source country in a destination as of 1990.[20]

Destination countries impose a variety of conditions in deciding which immigrants to admit, many of which involve the education level of immigrants. One indicator of the 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).[21] The European signatories of the Schengen Agreement have committed to abolish all border barriers, including temporary migration restrictions, 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 visitors from particular countries of origin, with the set of visa-waiver countries varying across destination countries. 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, the existing data do not permit one to characterize immigration policy very thoroughly in a manner that is comparable across destinations. As important as

[20] 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.

[21] 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.

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. One might imagine that source-country-specific characteristics could also affect international 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.[22]

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, equations (12), (13), and (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

[22] 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 variable was statistically insignificant in all specifications and its inclusion did not affect other results. The incidence of poverty in the source country may also affect credit constraints. While data on poverty headcounts are not available for all the countries in our sample, the share of agriculture in GDP tends to be highly correlated with poverty measures. The inclusion of the agriculture share of GDP also leaves our core results unchanged.

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.[23] 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.[24] 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.

Because the dependent variables have a log-odds metric, the magnitude of the regression coefficients does not have a particularly useful interpretation.[25] As a result, we focus in this section on the signs and significance levels of the coefficients. We discuss applications below that provide information about the quantitative effects of key variables on migration scale, selectivity, and sorting.

---

[23] 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 x 102 = 1530.

[24] 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.

[25] Based on equation (4), one might think that the coefficient on the earnings difference 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.

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 (these coefficients are not shown in order 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 similar.

In Table 4, all three wage coefficients are positive, as predicted by the theory. Furthermore, the coefficients from the selection and sorting regressions are quite similar and are both 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 that fixed costs are a function of observable characteristics of the source-destination pair. In the selection and sorting regressions, in contrast, fixed costs are differenced out. The difference in the wage coefficients between the scale and selection regressions suggests that 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.[26]

The regressions also include variables reflecting geographic, linguistic, social, and political relationships between source and destination countries. They show that language plays an important role in international migration. 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. Finally, the coefficients in the sorting regressions (columns (3) through (6)) show that English-speaking destinations attract higher-skilled immigrants than other destinations, on average.

The next variable is also language-related, indicating whether the source and destination countries have an official language in common. Its coefficients are positive

[26] In the LIS data, the U.S., the U.K, and Canada are first, fourth and fifth among destinations in terms of pre-tax wage differences and first, second, and third in terms of post-tax wage differences.

and significant, like the Anglophone-destination coefficients. Emigration is greater toward destinations that share a 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.

The next three variables capture differences in geography between the source and destination countries. 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 relative 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 long longitudinal distances) has no independent effect. The same two coefficients have different signs in the selection and sorting regressions. Emigrants to more distant destinations are more skilled than non-migrants, all else equal, but less skilled than emigrants to other destinations. The opposite is true of transoceanic emigrants.

History affects migration, too. Both short- and long-term 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 empirical results are consistent with these linkages disproportionately affecting migration of the less-skilled.

There is also an important role for our limited measures of immigration policy. 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 asylees and refugees may limit opportunities for more-skilled migrants to gain entry, producing a less skilled migrant inflow.[27] 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 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 layout of Table 5 is similar to Table 4. The dependent variables in Table 5 are the same as those in the corresponding columns of Table 4 and the units of observation are the same as well.

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,

[27] On asylee and refugee policy in Europe, see Hatton and Williamson (2004).

$\ln W_h^j - \ln W_s^j$. 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. As in Table 4, columns (1) through (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), (2), (3), and (5) and on post-tax data in columns (4) and (6). To focus on a case 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 equation (15) and the 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.[28] 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.[29] Thus, the sign pattern of the coefficients in Table 5 holds whether or not other regressors are included in the estimation.

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 that λ be positive. It also requires that for source-destination pairs where $\delta_h^3 - \delta_s^3 < 0$, migrants be negatively selected. In the data, we observe numerous cases where $\delta_h^3 - \delta_s^3 < 0$, but no negative selection. Inspection of

[28] 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 22).

[29] 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 – including the same regressors as in Table 4 either on their own or divided by the destination wage – the log wage variable enters with a negative sign in the scale and selection regressions.

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 joint assumptions that migrants maximize the log utility of net wages and that migration costs are proportional to wages.

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. Put differently, 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 relatively similar levels of labor productivity, sorting on log differences in wages (i.e., returns to skill) looks similar to sorting on level differences in wages. Indeed, the rank correlation between the log wage difference and the level wage difference across destination countries 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 widely differing levels of productivity.[30]

C. Robustness Checks

Tables 6 through 8 report results from a number of 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

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

presence of fixed migration costs. For the sorting regressions, we focus on specifications that include the post-tax wage differences. All of the estimates reported in these tables are taken from regressions that include all the variables reported in our baseline specifications, shown in columns (2), (4) and (6) of Table 4. Here 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.[31]

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 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. We conclude that the key results from the linear utility model are fairly robust to alternative wage measures.[32]

---

[31] The sample size is smaller here than for other LIS-based regressions because of missing PPP data for a few source countries.

[32] It would seem natural to treat the FO data as an instrument for the WDI data to deal with measurement error. To be a valid instrument, the measurement errors associated with the two different data sources would have to be uncorrelated with each other and with the true wage measures. Preliminary analysis showed that the covariance between the two measures exceeded the variance of the FO wage measure, which implies that the measurement errors are correlated with each other, with true wages, or both.

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.[33] 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

[33] Observations in which Ireland is the destination are dropped from these regressions because Ireland has no universities in the top 250.

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 considerably. 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 less-skilled individuals. Orrenius and Zavodny (2005) and McKenzie and Rapoport (2006) obtain similar findings for Mexican migration to the U.S. 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 two 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. Estimates based on the larger sample are similar to those from the smaller sample that includes only source countries with available wage data.

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.[34]

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

In this section we use the estimates from our regression analysis to shed light on different dimensions of international migration. These analyses provide insights into the scale, selectivity, and sorting of migrants. The first issue we address concerns fixed costs, which according to theory should play a role in determining the scale of migration. Despite the importance of fixed costs, there is little information on the magnitude of these costs in the literature. Our framework allows us to estimate fixed 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

[34] 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.

are given by $g_{sh}^{j} = x_{sh}\theta^{j}$, we obtain numerically identical estimates to those obtained by 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 U.S. dollars per year (the units in which wages are measured). We choose the Mexico-U.S. 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. 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.

Table 9 illustrates a number of points. The first is that fixed costs matter. 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.

Finally, 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 European countries. The U.S. 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.

Moving from the scale of emigration to the selectivity of emigrants, Table 10 decomposes migrant selectivity from different source countries by region and income level.[35] The first column gives mean selectivity of emigrants by source-region, income-level cell. Selectivity is the dependent variable in the selection regression, which is the difference in the log skill ratio between emigrants and non-migrants. Emigrants are positively selected on average from all source regions. Mean selectivity ranges from a high of 3.92 in low-income African countries to a low of 0.25 in the U.S. and Canada. It is generally lower in the high-income countries than in low-income countries.

The next three columns use the regression results reported in column (2) of Table 4 to decompose mean selectivity into components attributable to source-destination differences in skill-related wage differences, skill-varying migration costs, and a residual.[36] In Africa, migration costs account for about half of emigrant selectivity, with wage differences and the residual accounting for about one-quarter each. Emigrants from Africa appear to be much more positively selected than wage differences alone warrant.

[35] Low (high) income countries are those whose low-skill wage is below (above) the minimum value of this variable for the 15 OECD countries in the regression sample.
[36] The contribution of wage differences to selection is the mean across source countries of the destination-source difference in high and low-skill wages times the coefficient on wages in column (2) of Table 4; the contribution of migration costs is the mean across source countries of the sum of the regressors in column (2) of Table 4, each multiplied by its corresponding coefficient estimate; and the contribution of the residual is the mean across source countries of the residual for the regression in column (2) of Table 4.

In the other low-income countries, wage differences play a larger role, accounting for 37% of positive selection in Latin America and the Caribbean, 48% in low-income Asia, and 68% in Central and Eastern Europe. In high-income countries, migration costs contribute strongly to positive selectivity. Wage differences by themselves would actually contribute to negative selection in North America and high-income Asia and only modest positive selection in Western Europe, Australia, and New Zealand. Taking Tables 9 and 10 together, for most source countries, migration costs appear to play a large role in both how many individuals emigrate and which types emigrate.

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 11 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.[37] We seek to explain the immigrant skill gap, defined as the difference between the mean log skill ratio among immigrants in the U.S. 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. The decomposition explains the immigrant skills gap as a linear combination of the differences in mean values of the regressors, using the regression

[37] 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.

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.[38]

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 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.

The next three columns quantify the importance of distance. Contiguity has little effect. Longitudinal differences and distance have largely offsetting effects owing to the differing signs of their coefficients in Table 4. 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.

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

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 U.S. 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.

Our analysis also shows that language, history, and policy affect migration. English-speaking destinations draw higher-skilled immigrants than other destinations, whereas former colonial powers draw lower-skilled immigrants from their former colonies than from other source countries. 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: U.S. versus Europe." NBER Working Paper No. 9502.

Beine, Michael, Frederic Docquier, and Hillel Rapoport. 2001. "Brain Drain and Economic Growth: Theory and Evidence." *Journal of Development Economics*, 64(1): 275-289.

Beine, Michel, Frederic 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, Frederic 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.

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.

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 U.S. Immigration, 1971-1998." *Review of Economics and Statistics*, 89(2): 359-373.

Dahl, Gordon. 2002. "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets." *Econometrica* 70(6): 2367-2420.

Docquier, Frederic 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 McMillan, pp. 151-200.

Docquier, Frederic, 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: MacMillam.

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.

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. Luxembourg, periodic updating.

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-U.S. 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.

Psacharoupolos 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. 2005. "Trends in Total Migrant Stock: 2005 Revision." New York: Department of Economic and Social Affairs, United Nations.

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.

Figure 1: Share of emigrants and general population with tertiary education, 2000

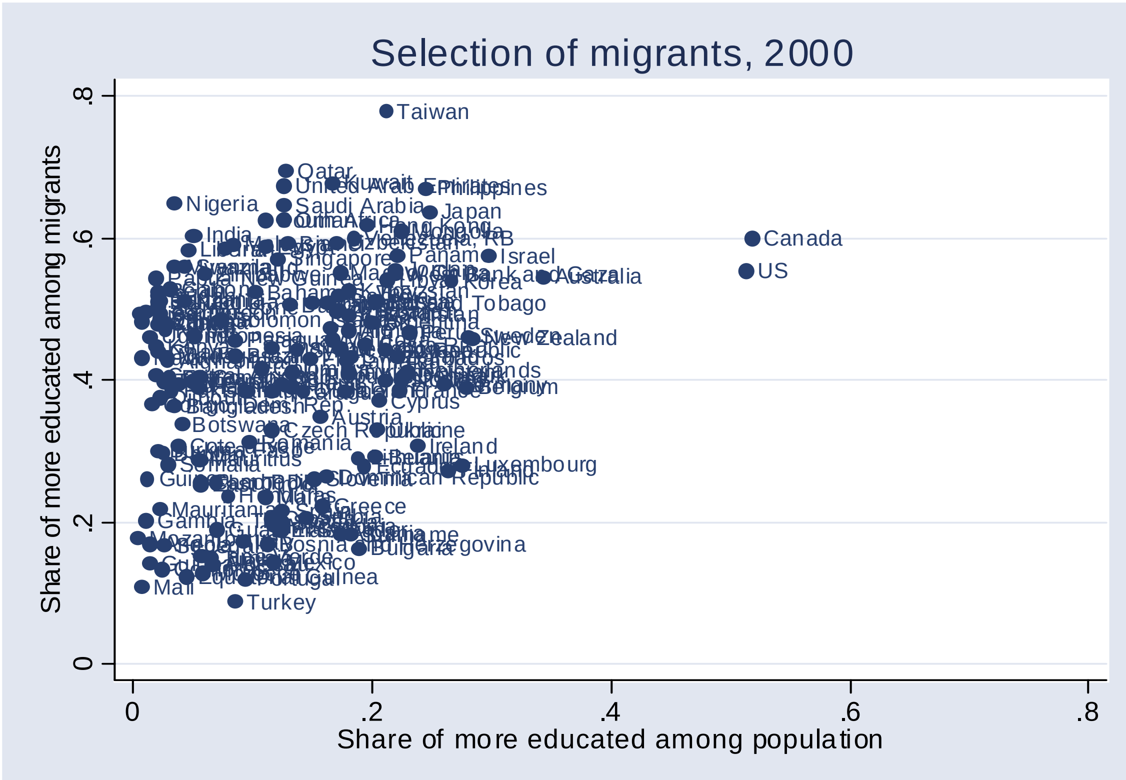

Figure 2: Emigration odds (primary and tertiary educated) by source country, 2000

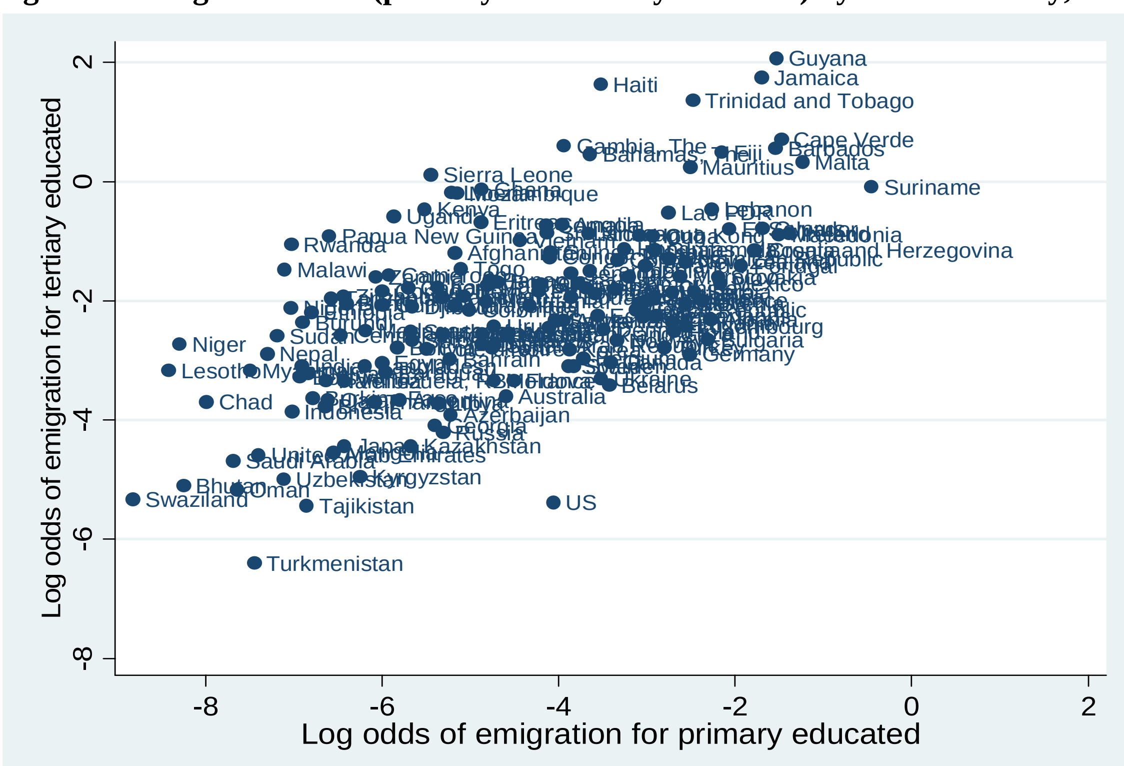

Table 1: Share of OECD immigrants by destination region and education, 2000

Destination Region	All	Education Group: 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 destination region and schooling group in 2000. North America includes Canada, Mexico and the United States; Australia and Oceania includes Australia, Japan, New Zealand, and South Korea; and Europe includes the other 24 OECD members (as of 2000).

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

	Destination region			
Source country	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

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

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		0.048
				(0.045)		(0.022)
Anglophone dest.	1.451	0.567	0.838	0.636	0.817	0.678
	(0.873)	(0.183)	(0.183)	(0.256)	(0.193)	(0.241)
Common language	0.648	1.268	0.355	0.352	0.331	0.332
	(0.293)	(0.248)	(0.137)	(0.139)	(0.125)	(0.124)
Contiguous	0.880	-0.384	-1.005	-1.007	-1.108	-1.097
	(0.401)	(0.373)	(0.229)	(0.237)	(0.230)	(0.240)
Longitude diff.	0.003	-0.009	0.004	0.004	0.005	0.005
	(0.004)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)
Log distance	-1.152	0.676	-0.245	-0.259	-0.273	-0.279
	(0.171)	(0.131)	(0.092)	(0.097)	(0.107)	(0.111)
LT colonial rel.	2.159	-0.711	-0.391	-0.445	-0.505	-0.550
	(0.411)	(0.193)	(0.176)	(0.161)	(0.150)	(0.137)
ST colonial rel.	2.641	-0.395	-0.129	-0.187	-0.195	-0.224
	(0.601)	(0.431)	(0.256)	(0.257)	(0.276)	(0.276)
Visa waiver	0.589	-0.299	0.335	0.364	0.440	0.471
	(0.314)	(0.135)	(0.164)	(0.172)	(0.200)	(0.203)
Schengen sig.	0.058	0.402	0.430	0.403	0.528	0.507
	(0.337)	(0.166)	(0.250)	(0.252)	(0.295)	(0.304)
Asylee share	-1.221	-2.512	-3.590	-3.635	-3.998	-4.007
	(3.698)	(0.818)	(0.901)	(0.709)	(0.929)	(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		5.338	
			(0.767)		(0.886)	
δ_h^3, post-tax				3.342		4.146
				(0.761)		(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:	Selection	Sorting
Variable	(1)	(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:	Selection	Sorting
Variable	(1)	(2)
$(W_h^3 - W_h^1)$, post-tax		0.047
		(0.024)
Observations		1202
R-squared		0.63
Clusters		
C. Freeman-Oostendorp wages		
Equation:	Selection	Sorting
Variable	(1)	(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:	Selection	Sorting
Variable	(1)	(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			0.078			0.082				
	(0.013)			(0.012)			(0.012)				
$(W_h^3 - W_h^1)$, post-tax		0.121	0.052		0.127	0.058		0.121	0.068	0.106	0.050
		(0.054)	(0.025)		(0.053)	(0.015)		(0.047)	(0.019)	(0.039)	(0.018)
Relative university quality				0.000	-0.002	-0.004					
				(0.001)	(0.002)	(0.002)					
Log emigrant stock 1990							-0.148	-0.039	-0.157		
							(0.031)	(0.055)	(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

A. Selection regressions; WDI wage measures

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	0.071	0.068	0.071	0.078	0.078	0.072	0.075	0.072	0.079	0.074	0.066	0.067	0.074	0.073
	(0.013)	(0.013)	(0.012)	(0.013)	(0.012)	(0.012)	(0.013)	(0.013)	(0.013)	(0.012)	(0.013)	(0.013)	(0.013)	(0.013)	(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

B. Sorting regressions; WDI post-tax wage measures

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	0.110	0.112	0.111	0.122	0.112	0.113	0.103	0.108	0.141	0.112	0.080	0.095	0.134	0.122
	(0.043)	(0.040)	(0.035)	(0.038)	(0.036)	(0.039)	(0.039)	(0.040)	(0.039)	(0.037)	(0.039)	(0.040)	(0.040)	(0.039)	(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

C. Sorting regressions; LIS post-tax wage measures

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	0.051	0.053	0.051	0.055	0.051	0.047	0.052	0.067	0.051	0.037	0.055	0.071
	(0.020)	(0.019)	(0.016)	(0.017)	(0.018)	(0.018)	(0.018)	(0.018)	(0.015)	(0.018)	(0.014)	(0.018)	(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		53.84	48.66	33.7	10.6
	21,375		8,910	7,998	49,954	715,825
China	80.28	56.8	90.83	88.57	87.73	55.6
	117,170	287,820	24,547	26,069	33,380	841,699
Colombia	89.88	69.4	72.84	71.72	71.43	34.67
	3,083	11,725	6,136	5,895	8,928	402,935
Dominican Republic	133.45	62.58	90.62	59.85	95.37	6.62
	38	3,225	238	3,188	350	527,520
El Salvador	50.52	26.82	83.06	82.21	89.85	-0.16
	6,314	27,780	328	332	302	619,185
Germany	34.69	19.94	14.8		13.76	21.83
	102,219	163,880	109,425		164,165	646,815
Guatemala	105.82	45.01	83.78	80.14	96.95	8.86
	157	8,880	346	448	212	341,590
Ireland	19.66	22.87	34.63	13.76	-28.28	16.55
	45,365	24,520	3,845	13,284	420,102	148,680
Italy	27.4	14.95	13.37	10.91	31.34	23.3
	216,316	312,185	371,714	456,000	86,876	461,085
Jamaica	76.79	-1.73	67.18	51.03	-9.31	-2.98
	680	103,265	299	872	124,313	449,795
Korea, Rep.	66.14	48.94	73.86	60.2	72.08	29.06
	25,160	50,860	6,164	12,226	7,434	676,640
Mexico	117.03	59.6	90.46	82.5	90.62	0
	870	24,795	3,064	4,029	3,558	6,374,825
Philippines	65.41	46.64	83.42	68.31	67.99	29.58
	78,105	191,615	4,767	12,539	34,782	1,163,555
Poland	45.63	26.38	27.95	5.21	42.89	29.83
	52,887	154,525	91,122	198,000	33,661	399,165
Portugal	46.13	9.32	-2.55	8.81	31.08	19.23
	13,329	143,145	538,106	113,216	26,006	187,645
Spain	66.39	63.67	15.77	27.11	42.88	50.85
	11,972	9,695	308,500	109,613	40,592	73,835
Turkey	63.17	68.03	40.1	7.95	53.27	60.36
	26,160	13,045	133,890	1,272,000	36,754	64,780
United Kingdom	12.39	11.87	23.02	17.29		36.99
	966,139	580,250	61,317	90,000		613,930
Vietnam	39.56	31.02	38.45	45.46	61.75	21.01
	128,666	127,590	58,570	43,105	19,137	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 emigrant selectivity, by income level and region of source country

Source countries	Mean selectivity	Mean contribution to selectivity of Wage difference	Migration costs	Residual
Low income countries in:				
Africa	3.918	0.955	2.014	0.950
Latin America & Car.	2.286	0.842	1.476	-0.033
Asia	1.982	0.949	1.352	-0.319
Cen. & E. Europe	1.265	0.862	0.864	-0.461
Australia/Oceania	--	--	--	--
High income countries in:				
Africa	--	--	--	--
North America	0.249	-0.210	1.393	-0.935
Asia	1.190	-0.038	1.333	-0.105
Western Europe	1.067	0.048	0.884	0.135
Australia & N. Zealand	0.891	0.193	1.479	-0.781
All countries	1.991	0.664	1.327	0.000

Notes: The table shows the mean value for the selection variable across source countries in column (1) (the log odds of migration for those with tertiary education relative to the log odds of migration for those with primary education), the mean value of the contribution of wage differences to selection in column (2), the mean value of the contribution of migration costs to selection in column (3), and the mean value of the residual in column (4). The relevant regression results are those in column (2) of Table 4. Low (high) income countries are those whose wage for low skill workers is below (above) the minimum value of this variable for OECD countries in the regression sample.

Table 11: Decomposition of the immigrant skills gap

Destination	Mean immigrant skills (1)	Immigrant skills gap (2)	Share of immigrant skills gap explained by: 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)	Share explained (14)
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
New 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
United Kingdom	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
United States	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.