THE ECONOMIC JOURNAL

The Economic Journal, **129** (*January*), 143–171 DOI: 10.1111/ecoj.12585
Advance Access Publication Date: 22 May 2018

SELF-SELECTION OF EMIGRANTS: THEORY AND EVIDENCE ON STOCHASTIC DOMINANCE IN OBSERVABLE AND UNOBSERVABLE CHARACTERISTICS*

George J. Borjas, Ilpo Kauppinen and Panu Poutvaara

The Roy model has more precise predictions about the self-selection of migrants than previously realised. The conditions shown to result in positive or negative selection in terms of expected earnings also imply a stochastic dominance relationship between the earnings distributions of migrants and non-migrants. We test these predictions using the Danish full population administrative data. We find strong evidence of positive self-selection of emigrants in terms of pre-emigration earnings and residual earnings: the income distribution for the migrants almost stochastically dominates the distribution for the non-migrants. Decomposing the self-selection in total earnings reveals that unobserved abilities play the dominant role.

A central finding in the economic literature on international migration is that emigrants are not randomly selected from the population of the source countries. The nature of the non-random selection affects the level and the distribution of welfare through two major channels. First, the skill distribution of migrants affects the wage structure in both sending and receiving countries (Borjas, 2003). A second effect takes place through the public sector. Immigration creates a fiscal surplus in the receiving country if and only if the net present value of the tax payments of immigrants exceeds the net present value of the costs they impose. Both the immigration of net recipients and the emigration of net payers pose a challenge to the public treasury (Wildasin, 1991; Sinn, 1997), while immigrants who are successfully integrated with the labour market can generate a substantial surplus (Dustmann and Frattini, 2014).

Beginning with Borjas (1987), there has been a great deal of interest in deriving and empirically testing models that predict how migrants differ from non-migrants. Many of these studies rely on an application of the Roy model of occupational self-selection. As long as skills are sufficiently transferable across countries, the sorting of persons across countries is mainly determined by international differences in the rate of return to skills. A country like the US would then attract high-skilled workers from more egalitarian

* Corresponding author: Panu Poutvaara, ifo Institute, Poschingerstrasse 5, 81679 Munich, Germany. Email: poutvaara@ifo.de.

We thank the Editor Frederic Vermeulen and two anonymous referees, Till Nikolka, Olli Ropponen, as well as participants at Norface Migration Network Conference, Journées Louis-André Gérard-Varet, EEA and CEMIR Junior Economist Workshop in 2013, the Alpine Population Conference, UCFS Workshop, CESifo ESP area conference and the VfS annual conference in 2015 and seminars at UC Irvine, ETH Zurich, Labour Institute for Economic Research, VATT, University of Linz and University of Salzburg for valuable comments and Carla Rhode for excellent research assistance. Financial support from Leibniz Association (SAW-2012-ifo-3) is gratefully acknowledged.

countries (i.e. countries offering relatively low rates of return to skills) and low-skilled workers from countries with greater income inequality (i.e. countries offering higher rates of return to skills). The evidence indeed suggests a negative cross-section correlation between the earnings of immigrants in the US and income inequality in the source countries.[1]

Although the existing literature on immigrant selection focuses mainly on the US context or on migration flows from poor to rich countries, there are also sizable migration flows between rich countries. According to the United Nations (2013), 21.9 million persons from EU15 countries now live outside their country of birth, with 42% of these migrants living in other EU15 countries and an additional 13% living in the US.[2]

This article examines the self-selection of emigrants from Denmark, one of the richest and most redistributive European welfare states. In 2013, over a quarter million Danes lived outside Denmark (corresponding to about 5% of the Danish-born population), with 50% of the migrants living in other EU15 countries and 13% in the US (United Nations, Department of Economic and Social Affairs, 2013). Emigration rates from Denmark are neither exceptionally low nor exceptionally large in a European comparison. In 2012, the emigration rate was in the age group 25–54 among the native-born 0.33% in Denmark, 0.18% in Germany, 1.19% in Ireland, 0.37% in the Netherlands, 0.13% in Spain and 0.29% in Sweden (Eurostat, 2016). Because the returns to skills in Denmark are relatively low, the canonical Roy model predicts that the emigrants should be positively selected in the sense that the expected earnings of the migrants exceed the expected earnings of the stayers.[3] However, there have been few systematic studies of the self-selection of migrants from a relatively egalitarian country to see whether this is indeed the case.[4]

Our theoretical analysis shows the same conditions that predict migrants are positively self-selected in the sense of a difference in expected incomes also predict that the income distribution of the migrants will first-order stochastically dominate the income distribution of the non-migrants.[5] The theory also distinguishes between selection in observable and selection in unobservable characteristics.

Our empirical analysis uses the Danish full population administrative data to study how migrants and non-migrants differ in their education, pre-emigration earnings and

[1] Related cross-country studies include Cobb-Clark (1993) and Bratsberg (1995). Grogger and Hanson (2011) examine the selection of migrants across a broad range of countries using an alternative theoretical framework where individuals maximise linear utility and migration is driven by absolute earnings differences between high and low-skilled workers.

[2] The EU15 countries were member states of the European Union prior to the expansion on May 1, 2004.

[3] For comparisons of gross wage premia from tertiary education across countries, see Boarini and Strauss (2010). A recent paper studying returns to cognitive skills is Hanushek *et al.* (2015). The study finds significant cross-country differences, with relatively low returns in Denmark and other Nordic countries.

[4] Studies of the selection of migrants across developed countries include Lundborg (1991), Pirttilä (2004), Kleven *et al.* (2014), Junge *et al.* (2014) and Gould and Moav (2016). Akcigit *et al.* (2016) use panel data on inventors from the US and European Patent Offices, and show that top tax rates strongly affect location choices of superstar inventors. Parey *et al.* (2017) study selection of university graduates who emigrate from Germany, using predicted earnings. Many studies also examine selection issues in a historical context; see Margo (1990), Ferrie (1996), Wegge (1999, 2002), Abramitzky and Braggion (2006) and Abramitzky *et al.* (2012).

[5] As those at the top of income distribution play a major role in job creation and innovations, self-selection of migrants at the top of income distribution can be expected to play a bigger role in the success of nations than the sheer numbers and their total income would suggest.

other observable characteristics. To shed light on the role of unobservable characteristics in the selection process, we investigate how migrants and non-migrants differ in terms of unobservable earnings ability, as measured by residuals from Mincerian earnings regressions. Our empirical results are in line with the predictions of the model: Danish emigrants are indeed positively self-selected both in terms of earnings and in terms of residuals from the wage regressions. Following our reframing of the canonical Roy framework in terms of the concept of stochastic dominance, our study specifically tests for whether the earnings distribution of the emigrants stochastically dominates that of the stayers (as would be predicted by the model). The evidence confirms this strong theoretical prediction over most of the support of the earnings distribution.

In our main analysis, we concentrate on long-term migrants who have spent at least five years abroad, but we also analyse separately migrants who have returned within five years (results in the technical appendix). We find empirical support for our stochastic dominance result for both groups, but more strongly among long-term migrants. This finding is in line with the Borjas and Bratsberg (1996) implication that return migration accentuates the type of selection of the initial immigrants.[6] Nonetheless, the differences between short-term and long-term migrants are small compared with the differences between migrants and non-migrants.

Our study is related to the flurry of papers that examine the selection of migrants from Mexico to the US. The pioneering analysis of Chiquiar and Hanson (2005) merged information from the US census on the characteristics of the Mexican migrants with information from the Mexican census on the characteristics of the Mexican non-migrants. Because the merged data did not report the earnings of migrants prior to the move, pre-migration earnings were predicted based on observable characteristics of the migrants. This 'counterfactual' empirical exercise suggested that Mexican emigrants were located in the medium-high range of the Mexican wage distribution. More recent studies by Fernández-Huertas Moraga (2011) and Kaestner and Malamud (2014) use survey data that report the actual pre-migration earnings and find evidence of negative selection. They also conclude that part of the negative selection can be traced to the unobservable characteristics that determine a migrant's earnings.[7]

In an important early contribution, Abramitzky (2009) analyses the effects of redistribution on internal migration in Israel. He uses a longitudinal data set to study how extensive redistribution by Israeli kibbutzim affects the self-selection of those who leave a kibbutz, and of those who enter a kibbutz. During the period of analysis, kibbutzim fully equalised their members' incomes, providing an ideal setting to test self-selection of those who leave an egalitarian community. Abramitzky (2009) finds strong support for the hypothesis that migrants' self-selection depends on returns to skills in the origin and in the destination. In terms of education, kibbutz-leavers are more skilled than those who stay and other rural migrants. Furthermore, those who had left a kibbutz had higher residual earnings than other rural migrants and non-migrants living in cities at a

[6] See Dustmann and Görlach (2016) for an excellent survey on temporary migration.

[7] There are two possible explanations for the negative selection in unobservable characteristics from Mexico: lower returns to unobservable skills in the destination country, or those with worse match quality being more likely to emigrate. Therefore, analysing migration from Mexico to the US does not allow distinguishing which mechanism is driving the results. Finding that the distribution of residuals of migrants stochastically dominates the distribution of residuals of non-migrants suggests that unobservable but internationally transferable skills play a bigger role than match quality.

later point. This testifies about positive self-selection of leavers in terms of unobservable characteristics, and is driven by less educated kibbutz-to-city movers.

The only other study that analyses self-selection of emigrants from a rich country in terms of residual earnings is Gould and Moav (2016), which complements our work. On the theory side, Gould and Moav (2016) make an important extension by allowing unobservable skills to have two components, one internationally applicable and one country-specific. They show that it is *a priori* unclear what shape self-selection with respect to residual income should take. In contrast, our analysis derives an unambiguous prediction of stochastic dominance with respect to total and residual income, but in a model that includes only one type of unobservable skills. On the empirical side, Gould and Moav (2016) find that the self-selection of Israeli men has an inverse U-shape in their residual earnings. We find that in Denmark, the probability of emigration increases strongly in residual earnings, with no inverse U-shape. Our model also helps to explain the negative selection in terms of unobservable characteristics that Fernández-Huertas Moraga (2011) and Kaestner and Malamud (2014) document among Mexican migrants to the US, and it rationalises the patterns that Abramitzky (2009) finds for intra-Israel migration.

The important role played by unobservable characteristics implies that constructing a counterfactual earnings distribution for the migrants based on observable characteristics can greatly bias the estimates. Our findings suggest that the use of such a counterfactual distribution will tend to understate the true selection in earnings, so that the selection implied by the counterfactual distribution is far weaker than the true selection – regardless of whether there is positive or negative selection. The numerical bias that results from using the counterfactual estimation is sizable in the Danish context: more than half of the difference between the expected earnings of migrants and non-migrants arises because of differences in unobserved characteristics.

The article is organised as follows. Section 1 sketches the economic theory underlying the analysis and derives theoretical predictions concerning the self-selection of emigrants, using the notion of stochastic dominance as a unifying concept. Section 2 introduces and describes the unique population data that we use and reports some summary statistics. Sections 3 and 4 present the main empirical findings in terms of observed and unobserved component of earnings. Section 5 evaluates the bias that results from predicting the pre-migration earnings of emigrants from the earnings distribution of non-migrants. Section 6 examines whether the selection of persons moving to other EU15 countries differs from the selection of migrants moving to countries where immigration restrictions come into play. Finally, Section 7 summarises the study and draws some lessons for future research.

1. Theoretical Framework

Previous literature on the self-selection of migrants has focused on the conditional expectations of earnings distributions among migrants and stayers. In this Section, we derive a novel result: the Roy model implies that under certain conditions, the earnings distribution of migrants' first-order stochastically dominates, or is stochastically dominated by, the earnings distribution of stayers. In a bivariate normal

framework, it turns out that the conditions required for stochastic dominance are identical to the conditions that determine the nature of self-selection in terms of expected earnings.

We also decompose self-selection into two components, one that is determined by differences in returns to observable skills between source and host country, and one that is determined by differences in returns to unobservable skills. The distinction between observable and unobservable skills, of course, depends on the empirical framework and on the data that is being used; observable skills include the variables explaining earnings that are included in the data, while the component of earnings that is left unexplained by the data is the unobservable skill component.

We take as our starting point the migration decision faced by potential migrants in a two-country framework, in line with Borjas (1987) and subsequent literature. Residents of the source country (country 0) consider migrating to the destination country (country 1), and the migration decision is assumed to be irreversible. To simplify the presentation, we focus on a single observed skill characteristic s and suppress the subscript that indexes a particular individual. For concreteness, the variable s can be thought of as giving the worker's years of educational attainment, but it includes all the characteristics affecting individual's income that are observed in a given set of data. Residents of the source country face the earnings distribution:

$$\log w_0 = \alpha_0 + r_0 s + \varepsilon_0, \tag{1}$$

where w_0 gives the wage in the source country; r_0 gives the rate of return to observable skills; and the random variable ε_0 measures individual-specific productivity shocks resulting from unobserved characteristics and is normally distributed with mean zero and variance σ_0^2. The distribution of observable skills in the source country's population is given by $s = \mu_s + \varepsilon_s$, where the random variable ε_s is also assumed to be normally distributed with mean zero and variance σ_S^2.

If the entire population of the source country were to migrate, this population would face the earnings distribution:

$$\log w_1 = \alpha_1 + r_1 s + \varepsilon_1, \tag{2}$$

where the random variable ε_1 is normally distributed with mean zero and variance σ_1^2. For analytical convenience, we assume that $\mathrm{Cov}(\varepsilon_0, \varepsilon_s) = \mathrm{Cov}(\varepsilon_1, \varepsilon_s) = 0$, so that the individual-specific unobserved productivity shocks (i.e. the 'residuals' from the regression line) are uncorrelated with observable characteristics.[8] The correlation coefficient between ε_0 and ε_1 equals ρ_{01}. The random variable ε_s is individual-specific and has the same value for the same individual in both countries, whereas ε_0 and ε_1 are both individual and country-specific.

Equations (1) and (2) completely describe the earnings opportunities available to persons born in the source country. Assume that the migration decision is determined by a comparison of earnings opportunities across countries net of migration costs C.

[8] A more realistic assumption would be that the correlation between observed and unobserved skills is positive. However, allowing for positive correlation does not change the qualitative predictions of the model.

Define the index function:

$$I = \log\left(\frac{w_1}{w_0 + C}\right) \approx [(\alpha_1 - \alpha_0) + (r_1 - r_0)\mu_s - \pi] + [(r_1\varepsilon_s + \varepsilon_1) - (r_0\varepsilon_s + \varepsilon_0)]$$

$$= \Delta\mu + (v_1 - v_0), \tag{3}$$

where π gives a 'time-equivalent' measure of migration costs ($\pi = C/w_0$). The cross-country difference in earnings net of the time-equivalent migration cost for an individual with average observed and unobserved characteristics is given by $\Delta\mu = [(\alpha_1 - \alpha_0) + (r_1 - r_0)\ \mu_s - \pi]$. The difference in earnings attributable to individual deviation from average characteristics is given by $(v_1 - v_0)$, where $v_i = (r_i\,\varepsilon_s + \varepsilon_i)$ for $i \in \{0, 1\}$. A person emigrates if and only if the index $I > 0$.

Migration costs vary among persons – but the sign of the correlation between costs (whether in dollars or in time-equivalent terms) and skills (both observed and unobserved) is ambiguous and difficult to determine. The heterogeneity in migration costs can be incorporated to the model by assuming that the distribution of the random variable π in the source country's population is given by $\pi = \mu_\pi + \varepsilon_\pi$, where μ_π is the mean level of migration costs in the population, and ε_π is a normally distributed random variable with mean zero and variance σ_π^2. However, Borjas (1987) and Chiquiar and Hanson (2005) show that time-equivalent migration costs do not play a role in the algorithm that determines the selection of emigrants if either those costs are constant (so that $\sigma_\pi^2 = 0$), or if the costs are uncorrelated with skills. For analytical convenience, we assume that time-equivalent migration costs are constant, so that $\pi = \mu_\pi$.[9] The outmigration rate from the source country is then given by

$$\Pr(I > 0) = \Pr(v^* > -\Delta\mu^*) = 1 - \Phi(-\Delta\mu^*), \tag{4}$$

where $v^* = (v_1 - v_0)/\sigma_v$ is a standard normal random variable; $\Delta\mu^* = \Delta\mu/\sigma_v$; $\sigma_v^2 = \mathrm{Var}(v_1 - v_0)$; and Φ is the standard normal distribution function.[10]

In addition to identifying the determinants of the outmigration rate in (4), the Roy model lets us examine which persons find it most worthwhile to leave the source country.[11] In the following, we examine the self-selection of emigrants along two dimensions: selection in terms of observable skills s and selection in terms of unobservable skills ε_0, which together combine into selection in terms of total productivity or earnings, as measured by $\log w_0$.

Let $F_M(z)$ and $F_N(z)$ represent the (cumulative) probability distributions of skills or earnings for migrants and non-migrants in the source country, respectively, where z denotes a particular measure of skills (e.g. observable or unobservable characteristics or income). By definition, the probability distribution of migrants $F_M(z)$ first-order

[9] If π were negatively correlated with skills, the negative correlation would tend to induce the more skilled to migrate, creating a positively selected migrant flow. This would strengthen positive self-selection, and weaken negative self-selection.

[10] It is straightforward to study (4) to confirm that the migration rate rises, when mean income in the source country falls, mean income in the host country rises, returns to observed skills in the source country fall, returns to observed skills in the host country rise, time-equivalent migration costs fall and when mean observed skills rise if $r_1 > r_0$ or fall if $r_1 < r_0$.

[11] Throughout the analysis, we assume that $\Delta\mu^*$ is constant. The migration flow is effectively assumed to be sufficiently small that there are no feedback effects on the labour markets of either the source or destination countries.

stochastically dominates that of stayers $F_N(z)$ if:

$$F_M(z) \leq F_N(z) \ \forall z, \tag{5}$$

and there is at least one value of z for which a strict inequality holds.[12] From now on, whenever we refer to stochastic dominance, we mean first-order stochastic dominance.

Equation (5) implies that a larger fraction of the migrants have skills above any threshold z^*. Put differently, for any level of skills z^*, the population described by the probability distribution F_M is more skilled because a larger fraction of the group exceeds that threshold. The migrants, in short, are positively selected. Negative selection, of course, would occur if the reverse was true and $F_N(z) \leq F_M(z)\ \forall z$, with a strict inequality holding for at least one value of z.

If the skill distribution of migrants stochastically dominates that of non-migrants, the stochastic dominance then also implies the typical definition of positive selection that is based on conditional expectations:

$$\mathrm{E}(z|I > 0) > \mathrm{E}(z|I \leq 0), \tag{6}$$

so that migrants, on average, are more skilled than stayers. Conversely, if the probability distribution of stayers stochastically dominates that of migrants, and there was negative selection, it would also follow that $\mathrm{E}(z \mid I > 0) < \mathrm{E}(z \mid I \leq 0)$. The converse, however, is not true for a general distribution: a claim of positive selection in expectations, as defined by (6), does not imply that the skill distribution of migrants stochastically dominates that of non-migrants.

To derive the stochastic ordering of the skill distributions of migrants and non-migrants, let $f(x, v)$ be a bivariate normal density function, with means (μ_x, μ_v), variances $(\sigma_x^2, \ \sigma_v^2)$ and correlation coefficient ρ. Furthermore, let the random variable v be truncated from below at point a and from above at point b. Arnold *et al.* (1993) show that the (marginal) moment generating function of the standardised random variable $(x - \mu_x)/\sigma_x$, given the truncation of v, is given by

$$m(t) = \left[\frac{\Phi(\beta - \rho t) - \Phi(\alpha - \rho t)}{\Phi(\beta) - \Phi(\alpha)}\right] e^{t^2/2}, \tag{7}$$

where $\alpha = (a - \mu_v)/\sigma_v$; and $\beta = (b - \mu_v)/\sigma_v$. In terms of the migration decision, the truncation in the random variable $v = v_1 - v_0$ in the sample of migrants is from below and implies that $\alpha = -\Delta\mu^* = k$, and $\beta = \infty$, where k is the truncation point. In the sample of stayers, the truncation in v is from above, and the truncation points are $\alpha = -\infty$ and $\beta = k$. By substituting these definitions into (7), it can be shown that the moment generating functions for the random variable giving the conditional distributions of skill characteristic x for migrants and stayers reduce to

$$m_F(t) = \left[\frac{1 - \Phi(k - \rho t)}{1 - \Phi(k)}\right] e^{t^2/2} \tag{8}$$

[12] An alternative and perhaps more intuitive definition of stochastic dominance is in terms of quantiles. Let $Q_M(P)$ and $Q_N(P)$ be the quantile functions of order P of the skill distributions of migrants and non-migrants. $F_M(z)$ stochastically dominates $F_N(z)$ if and only if $Q_M(P) \geq Q_N(P)$ for all $0 \leq P \leq 1$ and there is at least one value of P for which a strict inequality holds.

and

$$m_G(t) = \left[\frac{\Phi(k-\rho t)}{\Phi(k)}\right] e^{t^2/2}. \tag{9}$$

Consider any two distribution functions $F(z)$ and $G(z)$. Thistle (1993) shows that F will stochastically dominate G if and only if

$$m_F(-t) < m_G(-t), \forall t > 0, \tag{10}$$

where m_F is the moment generating function associated with distribution F; m_G is the moment generating function associated with G. The ranking of the moment generating functions in (10) implies we can determine the stochastic ranking of the two distributions by simply solving for the relevant correlation coefficient ρ, and comparing (8) and (9). Such a comparison implies that

$$\begin{aligned} F_M(z) < F_N(z), \quad \text{if} \quad \rho > 0, \\ F_M(z) > F_N(z), \quad \text{if} \quad \rho < 0. \end{aligned} \tag{11}$$

In other words, migrants are positively selected if $\rho > 0$, and are negatively selected otherwise. Consider initially the stochastic ranking in observable characteristics. The random variable $x = \varepsilon_s$, and the relevant correlation coefficient ρ is defined by

$$\rho = \text{Corr}(\varepsilon_s, \ v_1 - v_0) = \frac{r_0\sigma_s}{\sigma_v}\left(\frac{r_1}{r_0} - 1\right). \tag{12}$$

Equation (12) shows that the stochastic ordering of the distributions of observable skills of migrants and non-migrants depends only on international differences in the rate of return to observable skills. The skill distribution of migrants will stochastically dominate that of stayers when the rate of return to skills is higher abroad. Conversely, the skill distribution for non-migrants will stochastically dominate the distribution for migrants if the rate of return to observable skills is larger at home.

Consider next the stochastic ordering in the conditional distributions of unobservable skills ε_0. The relevant correlation for determining this type of selection is given by

$$\rho = \text{Corr} \ (\varepsilon_0, v_1 - v_0) = \frac{\sigma_0}{\sigma_v}\left(\rho_{01}\frac{\sigma_1}{\sigma_0} - 1\right). \tag{13}$$

It follows that the distribution of unobservable skills for migrants stochastically dominates that for non-migrants when $\rho_{01}(\sigma_1/\sigma_0) > 1$. Note that the necessary condition for positive selection has two components. First, the unobserved characteristics must be 'transferable' across countries, so that ρ_{01} is sufficiently high. Second, the residual variance in earnings is larger in the destination country than in the source country. The residual variances σ_0^2 and σ_1^2, of course, measure the 'price' of unobserved characteristics: the greater the rewards to unobserved skills, the larger the residual inequality in wages.[13] As long as unobserved characteristics are sufficiently transferable across

[13] This interpretation of the variances follows from the definition of the log wage distribution in the host country in terms of what the population of the source country would earn if the entire population migrated there. This definition effectively holds constant the distribution of skills.

countries, emigrants are positively selected when the rate of return to unobservable skills is higher in the destination.

Finally, consider the stochastic ranking in 'total' productivity. The earnings distribution in the source country given by (1) can be rewritten as

$$\log w_0 = (\alpha_0 + r_0\ \mu_s) + (r_0\ \varepsilon_s + \varepsilon_0) = (\alpha_0 + r_0\ \mu_s) + v_0, \tag{14}$$

where the normally distributed random variable v_0 has mean zero and variance σ^2_{v0}. The relevant correlation for determining the stochastic ranking of the earnings distributions of migrants and non-migrants is

$$\rho = \mathrm{Corr}(v_0, v_1 - v_0) = \frac{\sigma_{v0}}{\sigma_v}\left[\gamma\left(\frac{r_1}{r_0} - 1\right) + (1-\gamma)\left(\rho_{01}\frac{\sigma_1}{\sigma_0} - 1\right)\right], \tag{15}$$

where $\gamma = r_0^2\sigma_s^2/\sigma_{v0}^2$ and $1 - \gamma = \sigma_0^2/\sigma_{v0}^2$.

The sign of the correlation in (15), which determines the nature of the selection in pre-migration earnings, depends on the sign of a weighted average of the selection that occurs in observable and unobservable characteristics. Interestingly, the weight is the fraction of the variance in earnings that can be attributed to differences in observable and unobservable characteristics respectively. If there is positive (negative) selection in both 'primitive' types of skills, there will then be positive (negative) selection in pre-migration earnings. If, however, there are different types of selection in the two types of skills, the selection in each type is weighted by its importance in creating the variance of the earnings distribution. It is well known that observable characteristics (such as educational attainment) explain a relatively small fraction of the variance in earnings (perhaps less than a third). As a result, (15) implies that it is the selection in unobservables that is most likely to determine the nature of the selection in the pre-migration earnings of emigrants. This implication plays an important role in explaining why the evidence reported in Fernández-Huertas Moraga (2011) and Kaestner and Malamud (2014) conflicts with that of Chiquiar and Hanson (2005).

As mentioned earlier, the stochastic dominance results necessarily imply selection in terms of conditional expectations. In the case of bivariate normal distributions, the expectation of the earnings distribution of migrants $\mathrm{E}(\log w_0 | v^* > \Delta\mu^*)$ is given by

$$\mathrm{E}(\log w_0 | v^* > -\Delta\mu^*) = \alpha_0 + r_0\mu_s + \frac{r_0\sigma_s^2}{\sigma_v}\left(\frac{r_1}{r_0} - 1\right)\lambda(-\Delta\mu^*) + \frac{\sigma_0^2}{\sigma_v}\left(\rho_{01}\frac{\sigma_1}{\sigma_0} - 1\right)\lambda(-\Delta\mu^*), \tag{16}$$

where $\lambda(-\Delta\mu^*) = \phi\ (-\Delta\mu^*)/[1 - \Phi(-\Delta\mu^*)] > 0$, and ϕ is the density of the standard normal distribution. As can be seen by examining (16), the conditions that determine the self-selection in terms of expectations are the same as the conditions that determine the stochastic ordering of the skill distributions of migrants and non-migrants. In the normal distribution framework that underlies the canonical Roy model, stochastic dominance implies selection in expectations, and *vice versa*. In empirical applications, however, the prediction of stochastic dominance is likely to be much less robust than the predictions concerning expectations because testing for stochastic dominance will

require a more rigorous test than simply comparing the average incomes or skills of migrants and non-migrants. If one just compares the averages to find out how migrants are self-selected, the findings can be compatible with the predictions of the Roy model even if a large number of individuals in the data behave against the stochastic dominance predictions of the model. As a result, establishing an empirical pattern of stochastic dominance provides very strong evidence that differences in skill prices indeed play an important role in migration decisions.

2. Data

Our analysis uses administrative data for the entire Danish population from 1995 to 2010. The data are maintained and provided by Statistics Denmark and it derives from the administrative registers of governmental agencies that are merged using a unique social security number.[14] For each year between 1995 and 2004, we identified all Danish citizens aged 25–54 who lived in Denmark during the entire calendar year.[15] We restrict the analysis to persons who worked full time.[16] Migration decisions of part-time workers or of workers outside the labour force may be driven by different factors, and the observed income of these workers may not be indicative of their true earnings potential. The income variable for each year is constructed by adding the worker's annual gross labour income and positive values of freelance income.[17] We merged this information with data from the migration register for the years 1995 through 2010. The migration register reports the date of emigration and the country of destination. Even though it is possible for Danish citizens to emigrate without registering, we expect that the numbers of persons who do so is small as it is a legal requirement for Danish citizens to report emigration. Danish tax laws provide further incentives for migrants to register when they emigrate.

After identifying the population of interest, we determined for each person whether he or she emigrated from Denmark during the following calendar year. If we found that a particular person emigrated, we searched for the person in the migration register for subsequent years to determine if the migrant returned to Denmark at some point in the future, and recorded the date of possible return migration. The migration register includes near-complete information on return migration, as registration in Denmark is required for the return migrant to be eligible for income transfers and to be covered by national health insurance.

[14] All residents in Denmark are legally required to have a social security number. This number is necessary to many activities in daily life, including opening a bank account, receiving wages and salaries or social assistance, obtaining health care and enrolling in school.

[15] A person's age is measured as of January 1st the year after the reference year.

[16] The administrative data allow the calculation of a variable that measures the amount of 'work experience gained' during the calendar year. The maximum possible value for this variable is 1,000. We restrict our sample to workers who have a value of 900 or above, so that our sample roughly consists of persons who worked full time at least 90% of the year. In order to measure the work experience gained during a given year, we subtract the value from the previous year from the current value of the variable. Persons who had a missing value for work experience in either of the two years were dropped from the sample. Missing values in this variable typically indicate that the person spent time abroad.

[17] The information on earnings is taken from the tax records for each calendar year. This variable is considered to be of high quality by Statistics Denmark. Some persons also report negative values for freelance income. These negative values are likely to be due to losses arising from investments and do not reflect the productive characteristics of the individual.

Table 1

Summary Statistics

	Men			Women		
	Non-migrants	Nordic countries	Other destinations	Non-migrants	Nordic countries	Other destinations
Observations	6,450,665	2,104	5,219	5,163,129	993	2,443
Age						
Average	39.8	35.5	35.3	40.2	35.9	34.7
Median	40.0	33.0	33.0	40.0	34.0	33.0
Annual earnings in 2010 in euro						
Average	52,725	56,557	72,825	40,299	44,462	47,204
Median	46,675	49,646	61,283	37,976	41,235	43,109
Standardised annual earnings						
Average	1.0	1.1	1.4	1.0	1.1	1.2
Median	0.9	1.0	1.2	0.95	1.1	1.1
Education						
Comprehensive school	21.4	19.8	8.3	21.5	15.7	8.9
High school	3.2	7.8	8.6	3.1	6.9	8.9
Vocational school	49.8	43.5	30.3	41.8	36.5	30.8
Advanced vocational	5.6	5.7	6.6	4.9	5.1	7.8
Bachelor or equivalent	12.2	11.6	20.6	23.3	22.9	25.4
Master's or equivalent	7.3	10.6	23.9	5.1	12.3	17.6
Doctoral or equivalent	0.5	1.0	1.7	0.2	0.7	0.7

Notes. Standardised earnings are defined by the ratio of a worker's annual gross earnings to the mean gross earnings of workers of the same age and gender during the calendar year. The category 'advanced vocational' includes all the tertiary education programmes below the level of a Bachelor's programme or equivalent. Programmes on this level may be referred to for instance with such terms as community college education, advanced vocational training or associate degree.

We define a migrant as an individual who is found in one of the 1995–2004 cross-sections, who emigrates from Denmark during the following year to destinations outside Greenland or the Faroe Islands, and who stays abroad for at least five years.[18] Individuals who emigrated for less than five years are classified as short-term migrants, and the rest of the population is then classified as non-migrants. The analysis of both migrants and non-migrants is further restricted to only include Danish citizens who do not have an 'immigration background'.[19]

Table 1 reports summary statistics from the Danish administrative data. The panel data set contains yearly observations of over 6.4 million male and 5.1 million female non-migrants. The construction of the data implies that non-migrants appear in the data multiple times (potentially once in each cross-section). We were able to identify 7,323

[18] Greenland and the Faroe Islands are autonomous regions but still part of Denmark. We have excluded these destinations as many of these migrants could have originated in Greenland or the Faroe Islands. The exact duration requirements were 1,825 days or longer for long-term migrants. Among men (women), 36.51% (32.49%) of all migration spells lasted at least 1,825 days. Having stayed abroad for five years predicts longer migration spells: 72% of men and 71% of women who left Denmark in 1996 and were still abroad after five years were also abroad after 10 years.

[19] Statistics Denmark defines a person to have no immigrant background' if at least one of the parents was born in Denmark and the person is/was a Danish citizen. We searched the population registers from 1980 to 2010 for the parents of the persons in our sample, and if a parent was found he or she was required to be a Dane with no immigrant background as well.

male and 3,436 female migrants. By construction, these migrants are persons who we first observe residing in Denmark. Emigrants are younger than the non-migrants, regardless of gender. Despite being younger, the emigrants earned higher annual incomes in the year prior to the migration than the non-migrants.[20]

We construct a simple measure of 'standardised earnings' that adjusts for differences in age, gender and year effects. Standardised earnings are defined by the ratio of a worker's annual gross earnings to the mean gross earnings of workers of the same age and gender during the calendar year.[21] Table 1 shows that emigrants earn more than non-migrants in terms of standardised earnings. Table C3 in the online Appendix reports the number of emigrants moving to different destinations. The largest destinations for both men and women are two other Nordic countries, Sweden and Norway, as well as the US, the UK and Germany. These five countries account for 57% of all emigration. Table 1 also reports the education distributions for non-migrants and migrants. It is evident that the migrants tend to be more educated than the non-migrants, among both men and women. Among both men and women, the educational distribution of migrants to non-Nordic destinations stochastically dominates the distribution of non-migrants and migrants to Nordic destinations.[22]

In order to add time dimension, the evolution of the long-term emigration rate is presented in Figure B1(*a*) for men and in Figure B1(*b*) for women separately for the whole population and for those with higher education and without higher education. As we are looking at long-term migration, the emigration rates are small, but there is an upward trend. The rate is higher for men and for those with higher education. We also computed the difference between the average of the log standardised earnings, or a degree of selection, for migrants and non-migrants for each year from 1995 to 2004 for men and women separately. The results are reported in Figure B2(*a*,*b*) in the online Appendix. There is a downward trend in the difference for both men and women. The finding makes sense: when the migrants are positively self-selected and the emigration rate gets bigger, the average standardised earnings of migrants should get smaller. However, the variation across years is small, so that pooling the data is justified.

Table 2 reports the logit regressions for the binary decision whether to emigrate on other observables than earnings. The pattern in Table 1 is confirmed also when other controls are added: those with higher education are much more likely to emigrate. Furthermore, having partner or children reduces the likelihood of emigration. To summarise, the descriptive findings suggest a strong degree of positive selection – at least as measured by education and differences in the conditional means of earnings.

[20] The first two panels of Table C1 in the technical appendix provide detailed information on which percentage of the underlying population of non-migrants and those migrating for at least five years and being aged 18 or more is excluded by the age restrictions, and which percentage by the working time restriction. Once working time restriction is imposed, requiring positive gross income affects less than 0.04% of migrants. That a higher share of migrants is excluded by the working time requirement can be explained to a large extent by migrants being younger, and often migrating in connection with their studies, or just after completing their studies. The last two panels show that if the attention is restricted to those aged 30 to 54, the shares of migrants and non-migrants affected by the working time restriction are quite similar. Table C2 presents summary statistics for short-term migrants.

[21] Both migrants and non-migrants, as well as shorter-term migrants, are included in these calculations.

[22] Table C4 reports the education distributions for individuals who do not fulfil the full-time work requirement. Comparing Tables 1 and C4 shows that those not working full time are less educated. Importantly, migrants are positively self-selected in terms of education also among those not working full time.

Table 2

Logit Estimates of the Probability of Emigration, by Gender

	Men		Women	
	B	Or	B	Or
Married	−0.110**	0.896**	−0.191***	0.826***
	(0.04)	(0.03)	(0.05)	(0.04)
Children	−1.137***	0.321***	−1.232***	0.292***
	(0.05)	(0.02)	(0.07)	(0.02)
Married × Children	0.460***	1.585***	0.374***	1.453***
	(0.07)	(0.10)	(0.09)	(0.12)
High school	1.377***	3.961***	1.158***	3.184***
	(0.05)	(0.21)	(0.08)	(0.26)
Vocational school	0.186***	1.205***	0.159**	1.172**
	(0.04)	(0.05)	(0.06)	(0.07)
Advanced vocational	0.648***	1.911***	0.714***	2.043***
	(0.06)	(0.11)	(0.08)	(0.17)
Bachelor	1.097***	2.994***	0.581***	1.787***
	(0.04)	(0.13)	(0.06)	(0.11)
Master's	1.652***	5.220***	1.444***	4.239***
	(0.04)	(0.23)	(0.07)	(0.29)
PhD	1.723***	5.599***	1.655***	5.233***
	(0.10)	(0.58)	(0.21)	(1.11)
y1996	−0.032	0.969	−0.001	0.999
	(0.06)	(0.05)	(0.08)	(0.08)
y1997	0.002	1.002	−0.016	0.984
	(0.06)	(0.06)	(0.08)	(0.08)
y1998	−0.024	0.977	−0.001	0.999
	(0.06)	(0.05)	(0.08)	(0.08)
y1999	0.230***	1.259***	0.131	1.140
	(0.05)	(0.07)	(0.08)	(0.09)
y2000	0.260***	1.296***	0.238**	1.269**
	(0.06)	(0.08)	(0.09)	(0.11)
y2001	0.161**	1.175**	0.146	1.157
	(0.05)	(0.06)	(0.08)	(0.09)
y2002	0.208***	1.231***	0.046	1.047
	(0.05)	(0.06)	(0.08)	(0.08)
y2003	0.198***	1.220***	0.112	1.119
	(0.05)	(0.06)	(0.08)	(0.09)
y2004	0.246***	1.279***	0.178*	1.194*
	(0.05)	(0.07)	(0.08)	(0.09)
Constant	−6.700***	0.001***	−6.951***	0.001***
	(0.08)	(0.00)	(0.12)	(0.00)
Age fixed effects	Yes		Yes	
N	6,470,720		5,173,706	
Pseudo R^2	0.0540		0.0557	

Notes. The table reports logistic regression results. The dependent variable is a dummy that gets value one if the individual is a long-term migrant migrating to other Nordic countries or other destinations the following year. Individually clustered standard errors are in parentheses and the column labelled 'Or' reports odds ratios. Coefficients for the age fixed effects are not shown. The category 'advanced vocational' includes all the tertiary education programmes below the level of a Bachelor's programme or equivalent. Programmes on this level may be referred to for instance with such terms as community college education, advanced vocational training or associate degree. *p < 0.05, **p < 0.01 ***p < 0.001.

3. Selection in Pre-migration Earnings

This Section presents empirical evidence on the self-selection of emigrants from Denmark in terms of standardised pre-emigration earnings. Our theoretical framework predicts that the distribution of earnings for migrants should stochastically dominate that of non-migrants. As a result, our empirical analysis will mainly consist of comparing cumulative distributions of standardised earnings between migrants and non-migrants. An advantage of simply graphing and examining the cumulative distributions is that the analysis does not require any type of kernel density estimation, and that we do not need to impose any statistical assumptions or parametric structure on the data. We will also present in the technical appendix kernel density estimates of the earnings density functions as an alternative way of presenting the key insights. Finally, we will derive and report statistical tests to determine if the data support the theoretical prediction of stochastic dominance.

Figure 1(*a*) illustrates the cumulative earnings distributions for male migrants to Nordic countries, male migrants to destinations outside Nordic countries and for male non-migrants. The values of the standardised earnings are truncated at −2 and 2 to make the graphs more tractable.[23] The figure confirms that migrants were positively selected during the study period, but less strongly so to other Nordic countries.[24] This weaker selection may arise because the rate of return to skills in Nordic countries is relatively low when compared to that in other potential destinations.[25] Figure 1(*b*) presents corresponding evidence for women.[26] The main findings are qualitatively similar, but the positive selection seems weaker. This gender difference could be due to family ties attenuating women's self-selection; see Borjas and Bronars (1991) and Junge *et al.* (2014). The finding that self-selection to other Nordic countries is also positive may appear surprising, as the Roy model would not predict any clear pattern if the earnings distributions were identical between Denmark and other Nordic countries. However, Parey *et al.* (2017) show that earnings inequality among the high-skilled is lower in Denmark than in Sweden, Norway and Finland. Therefore, positive but weaker self-selection into other Nordic countries is consistent with what the Roy model predicts.[27] Figure B6(*a*) presents the corresponding kernel estimates of the density functions of the logarithm of standardised earnings for men, while Figure B6(*b*) presents the respective graphs for

[23] The truncation does not alter the results as only 0.07% of non-migrants, 0.19% migrants to other Nordic countries and 0.11% of migrants to other destinations lie below the lower truncation point. Correspondingly, 0.03% of non-migrants and 0.21% of migrants to destinations outside Nordic countries lie above the upper truncation point. There are no migrants to other Nordic countries above the upper truncation point. Furthermore, the statistical analysis of differences between cumulative distribution functions does not use truncation.

[24] Figures B3(*a*,*b*) show that the results are qualitatively similar if the analysis is restricted to the age group 30 to 54.

[25] Moreover, some Danes may live in southern Sweden but work in Denmark. This type of migration should decrease the estimated selection to Nordic countries as it is not related to returns to skills.

[26] For women, 0.06% of non-migrants lie below the lower truncation point and 0.00004% above the higher truncation point. There are no migrants lying below the lower or above the higher truncation point.

[27] Figures B4(*a*,*b*) show that long-term migrants emigrating outside the Nordic countries are somewhat more strongly positively selected than migrants who leave for less than five years, but the differences are small. Figures B5(*a*,*b*) show that the earnings distributions of short-term migrants to other Nordic countries do not differ from the earnings distributions of long-term migrants emigrating there.

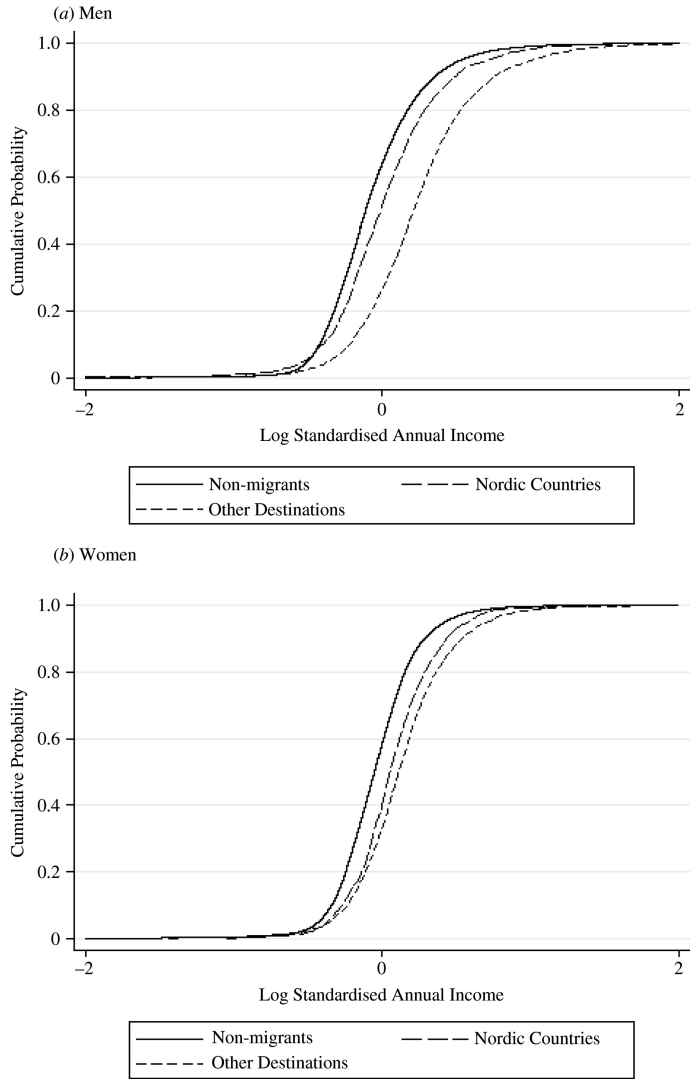

Fig. 1. *Distribution Functions of Standardised Annual Earnings for Migrants and Non-migrants. (a) Men (b) Women*

women.[28] The density functions again reveal the positive selection of migrants moving outside the Nordic countries, both for men and women.

As is evident from the figures, Kolmogorov–Smirnov tests comparing the earnings distributions for different groups reject the hypothesis that the underlying earnings distributions are the same at a highly significant level. Does the evidence statistically

[28] Following Leibbrandt *et al.* (2005) and Fernández-Huertas Moraga (2011), we use Silverman's reference bandwidth multiplied by 0.75 to prevent over-smoothing. The same bandwidth is used also in all the kernel density estimates reported in subsequent calculations.

support the theoretical prediction that the cumulative distribution function of migrants stochastically dominates that of non-migrants? Statistical tests for first-order stochastic dominance are highly sensitive to small changes in the underlying distributions, making it difficult to rank distributions in many empirical applications.[29] As noted by Davidson and Duclos (2013), it may be impossible to infer stochastic dominance over the full support of empirical distributions if the distributions are continuous in the tails, simply because there is not enough information in the tails for meaningful testing of any statistical hypothesis. It would then make sense to focus on testing stochastic dominance over a restricted range of the distribution. Based on these considerations, we apply an approach that characterises the range over which the value of the cumulative distribution function for non-migrants is statistically significantly larger than that of migrants. More formally, we seek for the longest interval $[\hat{w}^-, \hat{w}^+]$ for which the hypothesis

$$H_0 : \max_{w\epsilon[\hat{w}^-,\hat{w}^+]} [F_M(w) - F_N(w)] \geq 0, \tag{17}$$

can be rejected in favour of the alternative hypothesis of strong stochastic dominance[30]

$$H_1 : \max_{w\epsilon[\hat{w}^-,\hat{w}^+]} [F_M(w) - F_N(w)] < 0. \tag{18}$$

To perform the test, we estimate the difference between the cumulative distribution functions $\Delta[F(w)] = F_N(w) - F_M(w)$ with its sample value $\hat{\Delta}[F(w)]$ and calculate confidence intervals using tools that were introduced in Araar (2007) and Araar *et al.* (2009).[31] Davidson and Duclos (2000) show that the estimator $\hat{\Delta}[F(w)]$ is consistent and asymptotically normally distributed. Let $\hat{\sigma}(w)$ be the standard deviation of the estimator $\hat{\Delta}[F(w)]$. Furthermore, mark the significance level with θ and let $z(\theta)$ be the $100(1-\theta)$ percentile point of the standard normal distribution.[32] We can then define the lower bound for a one-sided confidence interval for $\Delta[F(w)]$ as[33]

$$\widehat{LB}_{\Delta(F(w))} = \hat{\Delta}[F(w)] - \hat{\sigma}(w)z(\theta). \tag{19}$$

We estimate the standard errors using a Taylor linearisation and allow for clustering at the individual level. We then implement the procedure by calculating the lower bounds of the confidence intervals for the estimate $\hat{\Delta}[F(w)]$ defined in (19).[34]

Although it is not clearly visible from Figure 1(*a*), the cumulative distribution functions of non-migrant men and men migrating to destinations outside the Nordic countries cross near the lower tails of the distributions. Figure B7(*a*) depicts $\hat{\Delta}[F(w)]$ and lower and upper bounds (calculated similar to the lower bounds) for a 95% confidence interval. The lower bound of the confidence interval is positive on most of the range

[29] This can lead to difficulties in empirical work, and less restrictive concepts such as restricted first-order stochastic dominance (Atkinson, 1987) and almost stochastic dominance (Leshno and Kevy, 2002) have been proposed.

[30] The alternative is strong dominance, as weak dominance cannot be separated statistically from the null hypothesis of non-dominance.

[31] The calculations are implemented using the DASP Stata module presented in Araar and Duclos (2013).

[32] The asymptotic variance of $\hat{\Delta}(w)$ is derived in Araar *et al.* (2009).

[33] Chow (1989) proved the theorem for the case of independent samples. Davidson and Duclos (2000) show that the results also extend to the case of paired incomes from the same population.

[34] Davidson and Duclos (2013) propose a bootstrap procedure that uses an empirical likelihood statistic. The procedure is asymptotically equivalent to the method we use. Given the large number of observations, the difference between the bootstrap and asymptotic test is likely to be very small. Moreover, as the asymptotic test is more conservative, our results err on the side of caution.

covering the supports of the distributions. Table C5 reports the shares of migrants and non-migrants whose earnings are outside the range over which the migrant distribution stochastically dominates at a 95% confidence level. Only 2.0% of the migrants and 3.4% of the non-migrants lie below the lower bound of the range where the lower bound of the confidence interval is positive, whereas the shares of migrants and non-migrants above the upper bound of the range are 0.1 and 0.0%. Put differently, the earnings of almost 98% of male migrants to destinations outside Nordic countries are on the range where the cumulative distribution function for non-migrants is statistically significantly above the function for migrants. Figure B7(*b*) depicts $\hat{\Delta}[F(w)]$ and the bounds for a 95% confidence interval for non-migrant women and women migrating to destinations outside Nordic countries. Only 2.8% of the migrants and 4.1% of the non-migrants have earnings below the range where the lower bound of the confidence interval is positive, and an even smaller 0.2% of the migrants and 0.0% of the non-migrants have earnings above this range. We interpret these findings as support for the stochastic dominance prediction for both men and women migrating outside Nordic countries.[35]

Additional support for our theory comes from Mexico. Our theory predicts that the earnings distribution of migrants from Mexico to the US should be stochastically dominated by the earnings distribution of non-migrants. Fernández-Huertas Moraga (2011) presents these distributions for men. Although he does not present confidence intervals as we do, the figures suggest a pattern that mirrors what we find for Denmark, reversing the curves for migrants and non-migrants. In Mexico, the wage distribution of non-migrants stochastically dominates that of migrants, apart from an overlap for a few percent at the bottom and converging at the top.

4. Selection in Unobserved Characteristics

In the previous Section, we documented the selection that characterises the migrants using the total pre-migration earnings (after adjusting for age and year). We now examine the component due to unobserved characteristics. In particular, we adjust for differences in educational attainment between migrants and non-migrants (as well as other observable variables) by running earnings regressions, and determine whether the distribution of the residuals differs between the two groups.

By construction, the residuals from a Mincerian wage regression reflect the part of earnings that is uncorrelated with the observed measures of skill. Obviously, the decomposition is somewhat arbitrary because it depends on the characteristics that are observed and can be included as regressors in the wage equation. Nevertheless, the study of emigrant selection in terms of wage residuals is important. First, selection in terms of unobservable characteristics sheds light on the importance of the quality of job matches relative to the skill component that is internationally transferable. The theory predicts that the nature of the selection in unobservable characteristics depends on the magnitude of the correlation coefficient measuring how the source and destination countries

[35] Figures B8(*a*,*b*) and the bottom panel of Table C5 present a corresponding analysis by comparing the cumulative distributions of migrants to other Nordic countries with that of non-migrants. Weaker positive self-selection into other Nordic countries can be rationalised by those countries having only slightly larger income differences among the high-skilled than Denmark (Parey *et al.*, 2017).

value these types of skills.[36] Second, the theory also suggests that the nature of the selection in pre-migration earnings depends on a weighted average of the selection that occurs in observable and unobservable characteristics, with the weights being the fraction of earnings variance attributable to each type of skill. Because observable characteristics play only a limited role in explaining the variance of earnings in the population, it is crucial to precisely delineate the nature of selection in unobservable characteristics.

Table 3 reports the Mincerian wage regressions that we use to calculate the residuals. The first two columns use annual earnings and include year dummies and age fixed effects. The last two columns use standardised earnings. The sample includes the whole population of prime aged full-time workers pooled over the entire 1995–2004 period. We also include the worker's marital status and number of children.[37]

Figure 2 presents the cumulative distributions of wage residuals for male (female) migrants to Nordic countries, male (female) migrants to destinations outside Nordic countries and male (female) non-migrants, when using non-standardised incomes. The values of the residuals are truncated at −2 and 2, a range that covers practically all of the population.[38] Among men, the cumulative distribution function of residuals for emigrants who moved outside the Nordic countries is located to the right of the cumulative distribution for migrants to Nordic countries, which in turn is located to the right of the cumulative distribution of the non-migrants. The visual evidence, therefore, provides a strong indication that migrants are positively selected in terms of unobserved characteristics. Figure 2(*b*) shows that female migrants are also positively selected in terms of wage residuals. As was the case when comparing the measure of pre-migration earnings in the previous Section, the selection in unobserved characteristics is less pronounced for women than for men. One explanation for this could be that men are typically primary earners in couples. Qualitative results are similar if using standardised earnings (Figure B9(a,*b*)), excluding marital status and dummy for having children (Figure B10(*a*,*b*)) or restricting the analysis to the age group 30 to 54 (Figure B11(*a*,*b*)).[39] We also performed Kolmogorov–Smirnov tests on the distributions of residuals for non-migrants and

[36] As long as this correlation is strongly positive (so that unobserved characteristics are easily transferable across countries), Danish emigrants would be positively selected in unobservables. After all, the payoff to these types of skills is likely to be greater in the destination countries. However, it could be argued that the correlation between the wage residuals in Denmark and abroad may be 'small'. For example, the residuals from the wage regression may be largely reflecting the quality of the existing job match in the Danish labour market, rather than measuring the worker's innate productivity. To the extent that the quality of the job match plays an important role in generating the residual, the correlation in this residual across countries would be expected to be weak (in fact, a pure random matching model would suggest that it would be zero). As a result, there would be negative selection in unobserved characteristics simply because Danish workers with good job matches (and hence high values of the residual) would not migrate.

[37] Table S6 shows that the results are quite similar when marital status and children are not included as controls, although the fit of the model is somewhat reduced.

[38] For men, 0.05% of non-migrants, 0.19% of migrants to other Nordic countries and 0.11% of migrants to other destinations lie below the lower truncation point. Correspondingly, 0.03% of non-migrants and 0.23% of migrants to destinations outside Nordic countries lie above the upper truncation point. There are no migrants to other Nordic countries above the upper truncation point. For women, 0.05% of non-migrants lie below the lower truncation point and 0.00% of non-migrants lie above the higher truncation point. There are no migrants lying below the lower or above the higher truncation point.

[39] A possible concern with the earnings regressions is that each individual is included in the data for every year he or she was living in Denmark and fulfilled the other conditions. At the same time, migrants are absent from the data from the migration year onwards. Table C7 shows the earnings regressions separately for 1995, 1999 and 2004 cross-sections to see if this has implications on the results. It has not. Cumulative distribution functions of standardised earnings (Figures B12–B14) and of residuals (Figures B15–B17) show that the selection stays qualitatively similar during the study period.

Table 3

Mincerian Earnings Regressions, by Gender

	(1) Men	(2) Women	(3) Men standardised	(4) Women standardised
	B	B	B	B
Married	0.068***	−0.016***	0.038***	−0.015***
	(0.00)	(0.00)	(0.00)	(0.00)
Children	0.025***	−0.048***	0.027***	−0.043***
	(0.00)	(0.00)	(0.00)	(0.00)
High school	0.224***	0.190***	0.232***	0.187***
	(0.00)	(0.00)	(0.00)	(0.00)
Vocational school	0.092***	0.089***	0.097***	0.086***
	(0.00)	(0.00)	(0.00)	(0.00)
Advanced vocational	0.186***	0.198***	0.190***	0.195***
	(0.00)	(0.00)	(0.00)	(0.00)
Bachelor	0.298***	0.225***	0.296***	0.222***
	(0.00)	(0.00)	(0.00)	(0.00)
Master's	0.498***	0.536***	0.496***	0.531***
	(0.00)	(0.00)	(0.00)	(0.00)
PhD	0.490***	0.622***	0.484***	0.620***
	(0.00)	(0.00)	(0.00)	(0.00)
1996	0.020***	0.017***		
	(0.00)	(0.00)		
1997	0.043***	0.041***		
	(0.00)	(0.00)		
1998	0.078***	0.083***		
	(0.00)	(0.00)		
1999	0.103***	0.112***		
	(0.00)	(0.00)		
2000	0.141***	0.143***		
	(0.00)	(0.00)		
2001	0.175***	0.175***		
	(0.00)	(0.00)		
2002	0.207***	0.210***		
	(0.00)	(0.00)		
2003	0.236***	0.235***		
	(0.00)	(0.00)		
2004	0.252***	0.258***		
	(0.00)	(0.00)		
Constant	12.131***	11.931***	−0.243***	−0.140***
	(0.00)	(0.00)	(0.00)	(0.00)
Age fixed effects	Yes	Yes	No	No
N	6,470,720	5,173,706	6,470,720	5,173,706
R^2	0.2597	0.3062	0.1652	0.1963

Notes. The table reports OLS results for earnings regressions. In models (1) and (2), the dependent variable is annual earnings in 2010 euro and in columns (3) and (4) standardised earnings. Standardised earnings are defined by the ratio of a worker's annual gross earnings to the mean gross earnings of workers of the same age and gender during the calendar year. Individually clustered standard errors are in parentheses. Coefficients for the age fixed effects are not shown. The category 'advanced vocational' includes all the tertiary education programmes below the level of a Bachelor's programme or equivalent. Programmes on this level may be referred to for instance with such terms as community college education, advanced vocational training or associate degree. *p < 0.05, **p < 0.01 ***p < 0.001.

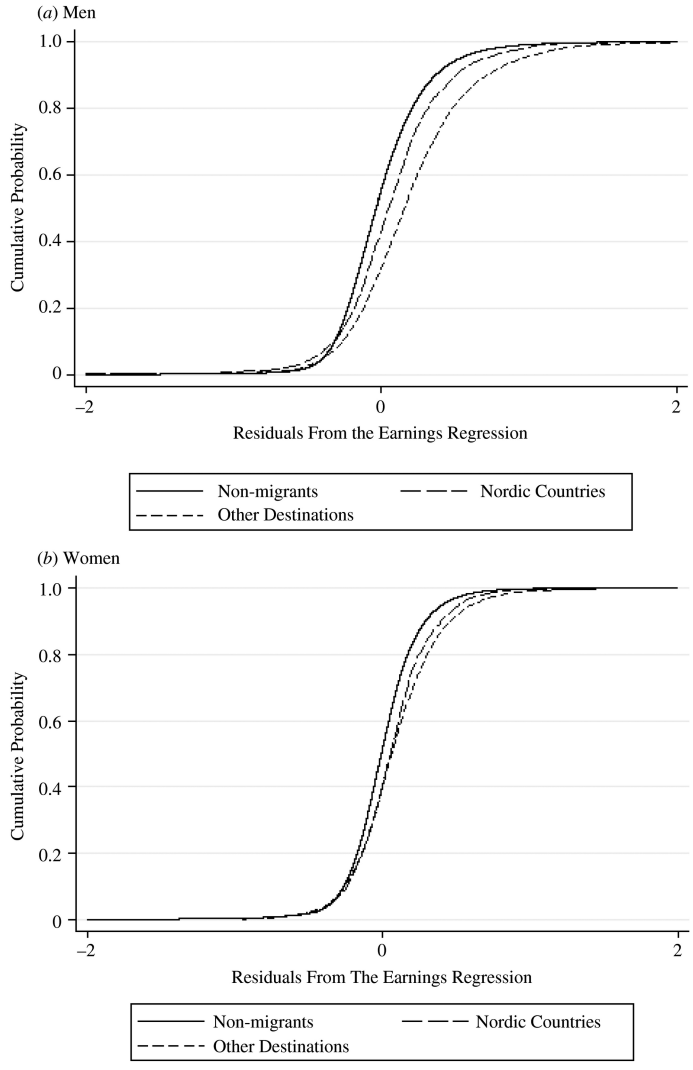

Fig. 2. *Distribution Functions of Residuals from Earnings Regression for Migrants and Non-migrants. (a) Men (b) Women*

migrants to other Nordic countries and for migrants to other destinations (separately for men and women). All the tests clearly rejected the null hypothesis, confirming that the distributions of residuals indeed differ among the groups.[40]

The evidence on the positive selection of migrants in unobserved characteristics obviously implies that the selection in pre-migration earnings documented in the previous Section cannot be attributed solely to the fact that migrants are more educated. Instead, we find that there is positive selection within education groups. This result also has implications on the interpretation of earnings regression residuals in general. The residuals

[40] The p-value for the test between women migrating to other Nordic countries and to other destinations was 0.015 and all the other p-values were 0.000, so that all tests clearly reject the hypothesis that the observations are drawn from the same distribution.

from wage regressions are sometimes interpreted as reflecting the value of the job match between the worker and the employer. If a high value for the residual only reflects a good match, we would then expect to find that workers with large residuals would be less likely to change jobs and less prone to migrate. Our findings clearly reject this interpretation.[41] It should be noted that our finding does not imply that job-specific matches would be unimportant; it only suggests that internationally transferable unobserved abilities are quantitatively clearly more important. Comparing results on the self-selection to other Nordic countries and the rest of the world suggests that search for a better job match to those who have a bad job match in Denmark is more pronounced among migrants to other Nordic countries.[42]

As in the previous Section, we also calculated the difference between the cumulative distribution functions with confidence intervals to determine whether empirical evidence supports the stochastic dominance prediction. The test results are summarised in Table C8. Figure B18(*a,b*) depicts $\hat{\Delta}[F(w)]$ and the lower and upper bounds for a 95% confidence interval for the comparison between non-migrant men (women) and men (women) migrating to destinations outside Nordic countries. The lower bound of the 95% confidence interval is positive on the range of residuals covering most of the support of the two distributions. Among men, 9.9% of the migrants and 15.2% of the non-migrants have wage residuals below the lower bound of this range, whereas the shares of migrants and non-migrants above the upper bound of the range are 0.1 and 0.0%. For women, 19.6% of migrants and 24.7% of non-migrants have earnings residuals below the lower bound of the range where the lower bound of the confidence interval is positive, and shares of migrants and non-migrants above the range are less than 1%.[43]

Figure 3 depicts the emigration rates outside other Nordic countries according to the deciles of residuals from the Mincerian wage regressions. Among men, the probability of emigration outside other Nordic countries first decreases slightly from 0.059% in the lowest decile of residuals to 0.038% in the second-lowest decile of residuals, and then increases monotonically in each decile, reaching 0.124% in the ninth decile and 0.232% in the highest decile. This contrasts sharply with the findings by Gould and Moav (2016) who studied 30 to 45-year old Israeli men in the 1995 census and found that the probability of emigration was inverse U-shaped in residuals. Among women, the probability of emigration from Denmark outside other Nordic countries first increases from 0.028% in the lowest decile of residuals to 0.035% in the second-lowest decile. It then decreases back to 0.028% in the third-lowest decile and then increases almost monotonically in each decile, reaching 0.062% in the ninth decile and 0.102% in the highest

[41] Finding negative selection of migrants in terms of residuals from a country with high returns to skills, like Mexico, does not allow distinguishing whether self-selection in residuals is driven by match quality or returns to unobservable skills.

[42] For this group, returns to unobserved productivity are not as important a criterion for self-selection as among migrants to the rest of the world, simply because differences in returns to skills between Denmark and other Nordic countries are smaller. As a result, the mechanism of searching for a better match quality is more pronounced.

[43] Figure B19(*a,b*) depicts $\hat{\Delta}[F(w)]$ and the bounds for a 95% confidence interval for non-migrant men (women) and men (women) migrating to other Nordic countries. The most notable result is that roughly 12% of both migrant and non-migrant men have residuals in the area where the upper bound of the confidence interval is negative.

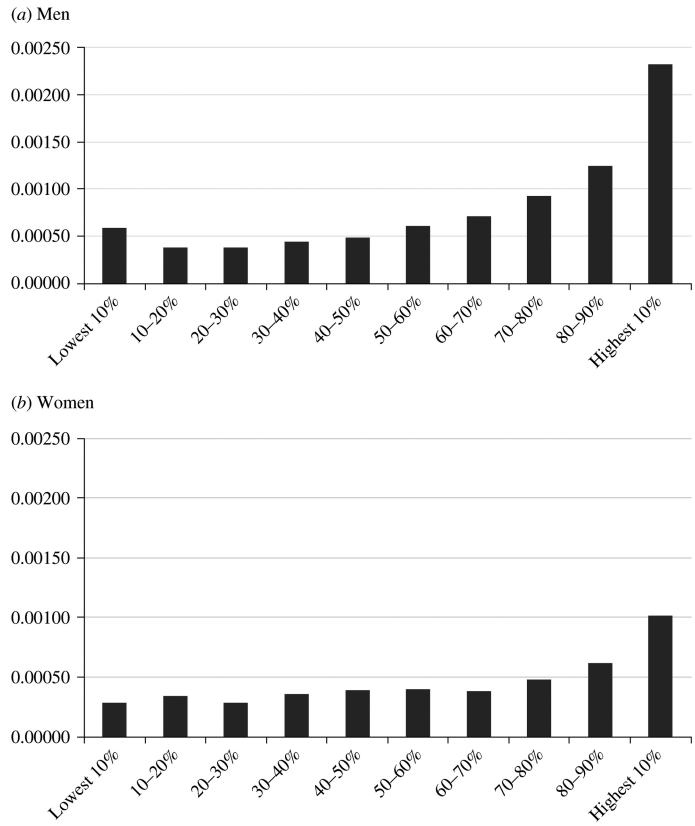

Fig. 3. *Annual Emigration Rate to Non-Nordic Destinations by Deciles of Residuals from Earnings Regressions. (a) Men (b) Women*

decile. Fernández-Huertas Moraga (2011) and Kaestner and Malamud (2014) find evidence of negative selection in terms of residuals when it comes to migration from Mexico to the US, in line with our theory. The stochastic dominance relationship among migrants from Denmark and Mexico and the inverse U-shape among emigrants from Israel suggests the need for further studies from other countries, to determine which pattern dominates globally.

We conclude by summarising the evidence as follows: there is strong positive selection in unobservable characteristics in the sample of migrants that moved outside the Nordic countries and weaker evidence of positive selection in the sample of migrants who moved to other Nordic countries.[44]

[44] Figures B20(*a*,*b*), B21(*a*,*b*) show that the differences in the self-selection in residuals between long-term migrants and short-term migrants are in line with the differences in terms of pre-migration earnings. Long-term migrants settling outside the Nordic countries are somewhat more strongly positively selected than short-term migrants, but the differences are small, while there are practically no differences between long-term migrants and short-term migrants emigrating to other Nordic countries.

5. Bias in Counterfactual Predictions

The fact that emigrants are self-selected in their unobserved characteristics implies that using the observable characteristics of migrants to predict their counterfactual earnings, had they chosen not to migrate, will lead to biased results. Due to data constraints, this is precisely the empirical exercise conducted by Chiquiar and Hanson (2005), who adopt the methodology introduced by DiNardo *et al.* (1996) and build a counterfactual wage density of what the Mexican immigrants would have earned in Mexico had they stayed. The actual wage density of Mexican 'stayers' is then compared to the counterfactual density for migrants. By construction, this approach ignores the role of unobservable characteristics in the estimation of the counterfactual wage distribution. A clear advantage of the Danish administrative data is that the earnings of emigrants can be observed before they emigrate, so there is no need to build a counterfactual density. The administrative data allows us to precisely measure the extent of the bias resulting from carrying out the counterfactual exercise in Chiquiar and Hanson (2005). In particular, we can contrast the predicted counterfactual wage distribution of migrants had they not moved to the actual wage distribution of migrants prior to their move. We carry out this exercise by precisely replicating the various steps in the Chiquiar–Hanson calculations, as depicted in the technical appendix. It is worth emphasising that this type of bias will arise not only in studies that examine the selection of migrants, but in any study that relies on observables to predict a counterfactual wage distribution.

Figure 4 presents the resulting counterfactual density functions of the logarithm of standardised earnings as well as the actual distributions for migrants and non-migrants.[45] The difference between the actual density for non-migrants and the counterfactual density for migrants reflects the part of self-selection that is due to observable characteristics. Similarly, the difference between the counterfactual and actual densities for migrants reflects the part of selection that is due to unobserved characteristics (i.e. all those variables that could not be included in the logit model).

One simple way of quantifying these distributional differences is to compute the averages of the various distributions. These calculations are reported in Table C9. Consider initially the results in the male sample. The difference between the mean of the actual distributions for migrants and non-migrants is 0.245 log points, but the difference between the counterfactual distribution and the distribution for non-migrants is 0.073. This implies that only about 30% of the positive selection in pre-migration earnings can be attributed to the observable characteristics included in the logit model, while about 70% is attributable to unobservable determinants of productivity. The calculations in the female sample yield a difference of 0.157 log points between the means of the actual distributions for migrants and non-migrants and a difference of 0.074 points between the counterfactual distribution and the distribution for non-migrants. As a result, observable and unobservable characteristics each account for about half of the positive self-selection in the pre-migration earnings of women.[46] The key lesson is clear: selection in

[45] To conduct the counterfactual analysis, we pool the sample of all migrants (regardless of whether they moved to Nordic countries or not).

[46] The component of self-selection that is due to unobservable characteristics plays a somewhat smaller role for women. One reason could be that women are more often tied migrants, and the migration decision may be mainly based on the skills of the spouse. The variance in income is also smaller for women, which also makes the selection both in terms of observable and unobservable characteristics weaker.

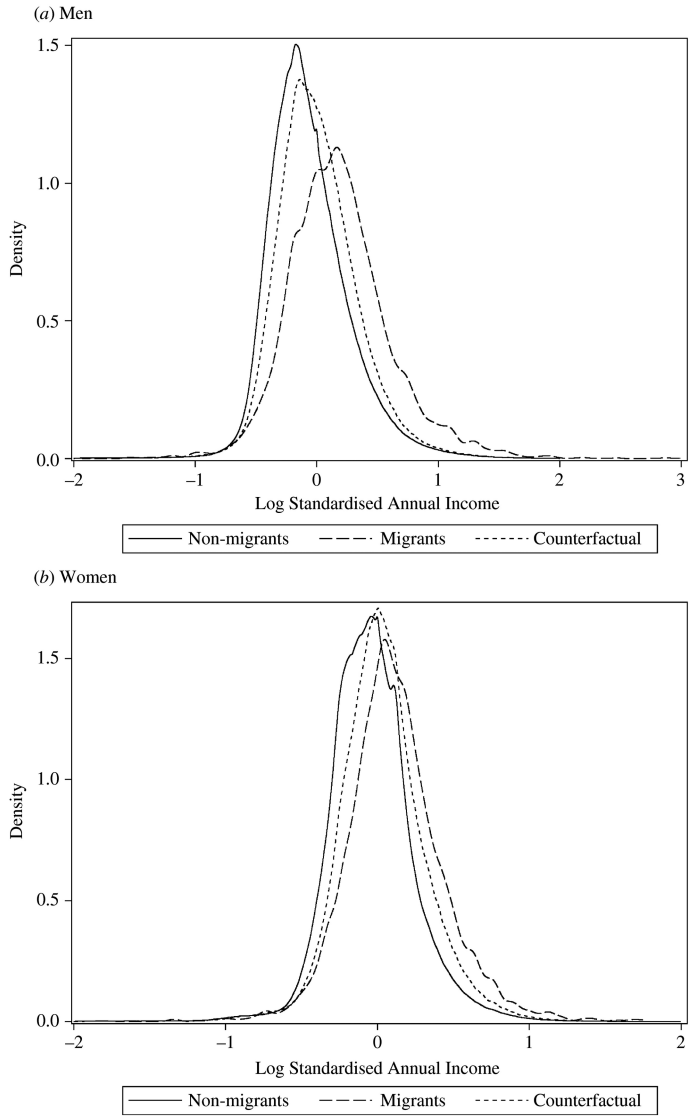

Fig. 4. *Counterfactual and Actual Densities of Standardised Gross Earnings. (a) Men (b) Women*

unobservable characteristics plays a crucial role in determining the skill composition of emigrants.

The distinct role of observables and unobservables in determining the selection in the pre-migration earnings is evident if we return to the Roy model and (16), which presents the conditional expectation $E(\log w_0 | v^* > -\Delta\mu^*)$. The nature of the selection in pre-migration earnings is given by the sum of the selection in observables and the selection in unobservables. Each of these selection terms has a weighting coefficient that represents

the variance in earnings attributable to observable characteristics ($r_0^2\sigma_s^2$) or to unobservable characteristics (σ_0^2). As noted earlier, observable characteristics explain a relatively small fraction of the variance in earnings. Put differently, (16) implies that it is the selection in unobservables that is most likely to determine the nature of the selection. To the extent that both selection in observables and unobservables work in the same direction, the counterfactual exercise will underestimate the true extent of positive selection in pre-migration earnings. Conversely, the counterfactual exercise will also attenuate the extent of 'true' negative selection if there is negative selection in both components of skills. Fernández-Huertas Moraga (2011) presents a corresponding analysis using survey data from Mexico and finds that counterfactual estimates greatly underestimate the extent of negative selection in the pre-migration earnings of Mexicans who move to the US. Put differently, the counterfactual exercise may lead to qualitatively right conclusions about the nature of the selection, but it may also greatly underestimate the true extent of either positive or negative selection.

6. Selection and Immigration Restrictions

As applied in the immigration literature, the Roy model focuses solely on the economic factors that motivate labour flows across international borders. The modelling typically ignores the fact that these flows occur within a policy framework where some receiving countries enact detailed restrictions specifying which potential migrants are admissible and which are not. We can use the administrative data from Denmark, combined with the unique political circumstances that guarantee free migration within Europe, to partially address the question of whether immigration policy affects selection all that much in the end. Specifically, we can subdivide the group of migrants who moved outside Nordic countries into two groups: those who moved to a country in the EU15 or to Switzerland, and those who moved to a country outside the EU15 and Switzerland. Movement of labour was unrestricted between Denmark and other EU15 countries and Switzerland in the period under study but was obviously restricted by immigration regulations to destinations outside the EU15, such as the US.

It turns out that these different immigration policies pursued by the EU15 and Switzerland and the rest of the world barely matter in determining the selection of Danish emigrants. Figure B22(*a*) depicts the cumulative distribution functions of the logarithm of standardised annual income for men and Figure B22(*b*) for women. It is evident that the distribution functions of standardised earnings are very similar for the two groups of migrants.[47] We also conducted the analysis using the wage residuals (not shown), and the distributions of residuals are also similar between the two groups. Given that Danish emigrants are typically highly educated, it is likely that they face relatively few restrictions. Immigration restrictions may play a more important role in explaining why emigrants from poorer and more unequal countries to rich countries are often relatively well educated, as established by Grogger and Hanson (2011).

[47] For women, a Kolmogorov–Smirnov test is not able to reject the null hypothesis that the observations for the two groups of migrants come from the same underlying distribution.

Our results also have tentative implications for the question of whether migration patterns reflect differences in rate of returns or migration costs. Although it is plausible that migration costs are higher when moving to other continents, our results suggest that such differences do not play a significant role in the sorting of emigrants between Europe and other continents, most notably North America.

7. Conclusion

This article shows that the Roy model has more dramatic predictions on the self-selection of emigrants than previously thought. The same conditions that have been shown to result in emigrants being positively (negatively) self-selected in terms of their average earnings actually imply that the earnings distribution of emigrants first-order stochastically dominates (or is first-order stochastically dominated by) the earnings distribution of non-migrants. Our theoretical analysis also distinguishes between selection in observable and selection in unobservable characteristics.

Our empirical analysis uses the Danish full population administrative data to analyse the self-selection of emigrants, in terms of education, earnings and unobservable ability, measured by residuals from Mincerian earnings regressions. The results are in line with the theory; the migrants are better educated and both pre-emigration earnings and wage regression residuals of migrants stochastically dominate those of non-migrants over most of the support of the distributions. Consider, for example, the case of full-time workers aged 25–54. For 98% of men and 97% of women who migrate outside other Nordic countries, the cumulative earnings distribution in the year before emigration stochastically dominates that of non-migrants with a 95% confidence interval. The difference between the cumulative distributions is not statistically significantly different in either direction for the remaining 2% to 3%.

Decomposing the self-selection in total earnings into self-selection in observable characteristics and self-selection in unobservable characteristics (as measured by residuals from Mincerian wage regressions), reveals that unobserved abilities play the dominant role. For men, about 70% of the positive self-selection in pre-migration earnings is attributable to unobservable determinants of productivity. For women, the fraction is about 50%. This suggests that relying on counterfactual distributions, based on observed characteristics, would strongly underestimate positive self-selection. This result complements the Fernández-Huertas Moraga (2011) finding that counterfactual estimates also greatly underestimate the extent of negative selection in the pre-migration earnings of Mexicans who move to the US. In short, the use of counterfactual earnings distributions based on observable characteristics greatly understate the true extent of selection in total earnings. Strong positive self-selection in residuals also suggests that unobserved abilities play a much bigger role in migration decisions than match quality, although it does not exclude a possibility that also match quality would be important. If most of the selection is on unobservables, then in settings with positive selection just focusing on the years of schooling would largely underestimate the brain drain of human capital that is lost by sending countries.

Our findings also have implications for immigration policies. Receiving countries can only base their admission policies on skill variables that are observed, whereas much of the selection of immigrants is 'hidden' in their unobserved characteristics. It can be expected that migrants will be self-selected in terms of unobserved characteristics even when admission restrictions are applied, and the self-selection among those fulfilling admission criteria can be expected to reflect relative skill prices. This raises a question about the effectiveness of point systems that are necessarily based on observable characteristics. The importance of relative skill prices is also supported by our separate analyses of self-selection of Danes migrating to the countries belonging to common European labour market (excluding other Nordic countries that have skill prices only marginally higher than Denmark) and not having any immigration restrictions, and the self-selection to the rest of the world. There is virtually no difference in the self-selection to these destination areas.

Harvard Kennedy School, NBER and IZA
VATT Institute for Economic Research
MU Munich, ifo Institute–Leibniz Institute for Economic Research at the University of Munich, CESifo, CReAM and IZA

Accepted: 15 September 2017

Additional Supporting Information may be found in the online version of this article:

Appendix A. Calculating the Counterfactual Earnings Distributions of Migrants.
Appendix B. Additional Figures.
Appendix C. Additional Tables.
Data S1.

References

Abramitzky, R. (2009). 'The effect of redistribution on migration: evidence from the Israeli kibbutz', *Journal of Public Economics*, vol. 93(3–4), pp. 498–511.
Abramitzky, R., Boustan, L. and Eriksson, K. (2012). 'Europe's tired, poor, huddled masses: self-selection and economic outcomes in the age of mass migration', *American Economic Review*, vol. 102(5), pp. 1832–56.
Abramitzky, R. and Braggion, F. (2006). 'Migration and human capital: self-selection of indentured servants to the Americas', *Journal of Economic History*, vol. 66(4), pp. 882–905.
Akcigit, U., Baslandze, S. and Stantcheva, S. (2016). 'Taxation and the international mobility of inventors', *American Economic Review*, vol. 106(10), pp. 2930–81.
Araar, A. (2007). 'Poverty, inequality and stochastic dominance, theory and practice: the case of Burkina Faso', Working Paper No. 2007-08, PEP-PMMA.
Araar, A. and Duclos, J.Y. (2013). User manual for Stata package DASP: version 2.3, Université Laval PEP, CIRPÉE and World Bank.
Araar, A., Duclos, J.Y., Audet, M. and Makdissi, P. (2009). 'Testing for pro-poorness of growth, with an application to Mexico', *Review of Income and Wealth*, vol. 55(4), pp. 853–81.
Arnold, B.C., Beaver, R., Groeneveld, R.A. and Meeker, W.Q. (1993). 'The nontruncated marginal of a truncated bivariate normal distribution', *Psychometrika*, vol. 58(3), pp. 471–88.
Atkinson, A.B. (1987). 'On the measurement of poverty', *Econometrica*, vol. 55(4), pp. 749–64.
Boarini, R. and Strauss, H. (2010). 'What is the private return to tertiary education? New evidence from 21 OECD countries', *OECD Journal: Economic Studies*, vol. 2010(1), pp. 1–25.
Borjas, G.J. (1987). 'Self-selection and the earnings of immigrants', *American Economic Review*, vol. 77(4), pp. 531–53.

Borjas, G.J. (2003). 'The labor demand curve is downward sloping: reexamining the impact of immigration on the labor market', *Quarterly Journal of Economics*, vol. 118(4), pp. 1335–74.
Borjas, G.J. and Bratsberg, B. (1996). 'Who leaves? The outmigration of the foreign-born', *Review of Economics and Statistics*, vol. 78(1), pp. 165–76.
Borjas, G.J. and Bronars, S.G. (1991). 'Immigration and the family', *Journal of Labor Economics*, vol. 9(2), pp. 123–48.
Bratsberg, B. (1995). 'The incidence of non-return among foreign students in the United States', *Economics of Education Review*, vol. 14(4), pp. 373–84.
Chiquiar, D. and Hanson, G.H. (2005). 'International migration, self-selection, and the distribution of wages: evidence from Mexico and the United States', *Journal of Political Economy*, vol. 113(2), pp. 239–81.
Chow, K.V. (1989). 'Statistical inference for stochastic dominance: a distribution free approach', Ph.D. Thesis, University of Alabama.
Cobb-Clark, D. (1993). 'Immigrant selectivity and wages: the evidence for women', *American Economic Review*, vol. 83(4), pp. 986–93.
Davidson, R. and Duclos, J.Y. (2000). 'Statistical inference for stochastic dominance and for the measurement of poverty and inequality', *Econometrica*, vol. 68(6), pp. 1435–64.
Davidson, R. and Duclos, J.Y. (2013). 'Testing for restricted stochastic dominance', *Econometric Reviews*, vol. 32(1), pp. 84–125.
DiNardo, J., Fortin, N.M. and Lemieux, T. (1996). 'Labor market institutions and the distribution of wages, 1973–1992: a semiparametric approach', *Econometrica*, vol. 64(5), pp. 1001–44.
Dustmann, C. and Frattini, T. (2014). 'The fiscal effects of immigration to the UK', Economic Journal, vol. 124(580), pp. F593–643.
Dustmann, C. and Görlach, J.-S. (2016). 'The economics of temporary migrations', *Journal of Economic Literature*, vol. 54(1), pp. 98–136.
Eurostat (2016). 'Emigration by five year age group, sex and country of birth', (migr_emi4ctb) available at: http://ec.europa.eu/eurostat/en/data/database (last accessed: 4 January 2016).
Fernández-Huertas Moraga, J. (2011). 'New evidence on emigrant selection', *Review of Economics and Statistics*, vol. 93(1), pp. 72–96.
Ferrie, J. (1996). 'A new sample of males linked from the public use micro sample of the 1850 US federal census of population to the 1860 US federal census manuscript schedules', *Historical Methods*, vol. 29(4), pp. 141–56.
Gould, E.D. and Moav, O. (2016). 'Does high inequality attract high skilled immigrants?', *Economic Journal*, vol. 126(593), pp. 1055–91.
Grogger, J. and Hanson, G.H. (2011). 'Income maximization and the selection and sorting of international migrants', *Journal of Development Economics*, vol. 95(1), pp. 42–57.
Hanushek, E.A., Schwerdt, G., Wiederhold, S. and Woessmann, L. (2015). 'Returns to skills around the world: evidence from PIAAC', *European Economic Review*, vol. 73(C), pp. 103–30.
Junge, M., Munk, M.D. and Poutvaara, P. (2014). 'International migration of couples', Working Paper No. 4927, CESifo.
Kaestner, R. and Malamud, O. (2014). 'Self-selection and international migration: new evidence from Mexico', *The Review of Economics and Statistics*, vol. 96(1), pp. 78–91.
Kleven, H.J., Landais, C., Saez, E. and Schultz, E. (2014). 'Migration and wage effects of taxing top earners: evidence of the foreigners' tax scheme in Denmark', *Quarterly Journal of Economics*, vol. 129(1), pp. 333–78.
Leibbrandt, M., Levinsohn, J. and McCrary, J. (2005). 'Incomes in South Africa since the fall of apartheid', Working Paper No. 11384, NBER.
Leshno, M. and Kevy, H. (2002). 'Preferred by all and preferred by most decision makers: almost stochastic dominance', *Management Science*, vol. 48(8), pp. 1074–85.
Lundborg, P. (1991). 'Determinants of migration in the Nordic labor market', *Scandinavian Journal of Economics*, vol. 93(3), pp. 363–75.
Margo, R.A. (1990). *Race and Schooling in the South, 1880–1950: An Economic History*, Chicago, IL: University of Chicago Press.
Parey, M., Ruhose, J., Waldinger, F. and Netz, N. (2017). 'The selection of high-skilled emigrants', *Review of Economics and Statistics*, vol. 99(5), pp. 776–92.
Pirttilä, J. (2004). 'Is international labour mobility a threat to the welfare state? Evidence from Finland in the 1990's', *Finnish Economic Papers*, vol. 17(1), pp. 18–34.
Sinn, H.-W. (1997). 'The selection principle and market failure in systems competition', *Journal of Public Economics*, vol. 66(2), pp. 247–74.
Thistle, P.D. (1993). 'Negative moments, risk aversion and stochastic dominance', *Journal of Financial and Quantitative Analysis*, vol. 28(2), pp. 301–11.

United Nations, Department of Economic and Social Affairs (2013). 'Trends in international migrant stock: migrants by destination and origin', (United Nations database, POP/DB/MIG/Stock/Rev.2013).
Wegge, S.A. (1999). 'To part or not to part: emigration and inheritance institutions in nineteenth-century Hesse-Cassel', *Explorations in Economic History*, vol. 36(1), pp. 30–55.
Wegge, S.A. (2002). 'Occupational self-selection of European emigrants: evidence from nineteenth-century Hesse-Cassel', *European Review of Economic History*, vol. 6(3), pp. 365–94.
Wildasin, D.E. (1991). 'Income redistribution in a common labor market', *American Economic Review*, vol. 81(4), pp. 757–74.