Let's be selective about migrant self-selection*

Costanza Biavaschi[†] Benjamin Elsner[‡]

February 19, 2014

Abstract

Migrants typically differ from the average population in their home country. While the causes of this difference — known as self-selection — have been documented for many countries, in this paper we turn to its consequences. Using a combination of non-parametric estimation and calibrated simulation, we quantify the impact of migrant self-selection on GDP per capita in both sending and receiving countries. Two episodes of mass migration serve as examples: the migration from Norway to the US in the 1880s and from Mexico to the US in the 2000s. We first show that Norwegians were mildly positively and Mexicans negatively selected from their home country population. In a simulation exercise, we then compare the economy under selective migration with a counterfactual in which the same number of migrants are neutrally selected. In both periods, self-selection had virtually no effect in the US. In the sending countries, the impact was small in Norway but substantial in Mexico: it reduced Norwegian GDP per capita by 0.26%, while it increased Mexican GDP per capita by 1.1%. The results suggest that researchers should be careful when claiming that migrant self-selection has large welfare implications.

*We would like to thank Samuel Bazzi, Anthony Edo, Alan Fernihough, Jesús Fernández-Huertas Moraga, Boris Hirsch, Volker Lindenthal, Luca Marchiori, Julia Matz, Sebastian Siegloch, as well as participants at ETSG 2013, CESifo, IOS Regensburg, University of Münster, OECD Paris, and FERDI for helpful comments.

[†]Institute for the Study of Labor (IZA). Address: Schaumburg-Lippe-Str. 5-9, 53113 Bonn, Germany. biavaschi@iza.org.

[‡]Corresponding author. Institute for the Study of Labor (IZA). Address: Schaumburg-Lippe-Str. 5-9, 53113 Bonn, Germany. elsner@iza.org, www.benjaminelsner.com.

1 Introduction

Migrant self-selection — the question who migrates and who doesn't — is a fundamental issue in the economics of migration. The literature has found a significant degree of self-selection for migrants from virtually all major sending countries. Nonetheless, while its causes are well-understood, and while many studies are motivated by the potential welfare effects, the consequences of self-selection are far from clear. It may well be that emigrants are younger, more educated, and more motivated than the average person in their home country, but does this difference really matter for the sending or receiving countries? We address this question by estimating the welfare impact of migrant self-selection using the mass migration waves from Norway in the 1880s and from Mexico in the 2000s.

Both episodes are examples of the two largest migration waves in the history of the US: the mass migration from Europe in the 19th century and the Mexican migration to the US since the 1980s. Despite being 120 years apart, both episodes are more comparable than they might initially appear. 9% of the population left Norway and Mexico and settled in the US. Moreover, in both cases the GDP per capita was around 30% of US GDP at the given time. The main difference between both episodes lies in the selection pattern of emigrants. As we show, Norwegian emigrants were mildly positively selected, meaning that they were more skilled than the average Norwegian. Mexican emigrants, in contrast, were less skilled than the average Mexican, and thus negatively selected.

To quantify the aggregate impact of migrant self-selection, we combine non-parametric estimation with a calibrated simulation exercise. For both sending countries, we use panel data that provides information on migrants before and after migration. Using earnings before migration as a measure for skills, we estimate the degree of self-selection as the difference between the skill distributions of migrants and the entire population of the sending country. To obtain the panel data for Norway, we match newly available historical census records based on name and birth year. For Mexico, we use the ENET survey, which contains all the relevant information.

In a second step, we feed the estimated distributions into a calibrated general equilibrium model based on Yeaple (2005) and Iranzo & Peri (2009), and compare GDP per capita under the current selection pattern and under neutral selection. To understand the underlying thought experiment, consider the migration of 10 million negatively selected Mexicans to the US. We first repatriate all these 10 million migrants, then randomly draw 10 million new migrants from the total population, and send them back to the US. Because the number of migrants is kept constant, the resulting effect is purely driven by self-selection.

Within the model, self-selection affects GDP per capita through two channels: the labor market channel and the productivity channel. A change in migrants' skills changes the nominal wage structure, as it increases the labor market competition for some skill levels, while decreasing it for others. Quantitatively more important than the nominal wage channel is the productivity channel. If migrant self-selection makes the workforce more productive on average, aggregate prices decrease, which is equivalent to an increase in per-capita GDP.

Our results demonstrate that migrant self-selection can — but does not necessarily have to — have a significant aggregate impact. Indeed, it only matters if both the size of the migration flow *and* the degree of self-selection are sufficiently large. In both periods, we find virtually no effect on the US economy. The influx of 180,000 Norwegians was simply too small to have any impact in the US. While the influx of 10 million Mexicans in the 2000s increased the US population by 4%, the effect of selection on GDP per capita only amounted to +0.28%. The reason for this small effect is the low degree of skill transferability of Mexicans in the US. Mexicans are so heavily concentrated at the lower end of the US skill distribution that even a substantial change in their skill selection does not result in a large aggregate impact. Due to the low degree of selection, the aggregate impact on Norwegian GDP is equally small; positive selection reduces Norwegian GDP in 1880 by 0.26%. In Mexico, which had a large emigration wave with a significant negative selection, the effect is considerably larger. Because of negative selection, Mexican per-capita GDP is 1.1% higher than it would be if migrants had the same skills as the average Mexican. While this effect might appear small at first, additional simulations show that it is as large as the difference in GDP per capita between zero migration and the current level of migration.

With its focus on the consequences of self-selection, this paper offers a new perspective on the literature on migrant self-selection.[1] In particular, it complements previous studies on the causes of self-selection from Norway (Abramitzky *et al.*, 2012) and Mexico (Chiquiar & Hanson, 2005; Fernández-Huertas Moraga, 2011, 2013; Ambrosini & Peri, 2012; Kaestner & Malamud, 2013). In both cases, the literature has shown that migrants significantly differ from the total population. We first confirm these results, before proceeding to demonstrate their implications for the sending and receiving countries. While many studies on the causes of self-selection are motivated by its potential welfare impacts, our paper shows that self-selection has no significant aggregate impact in 3 out of 4 cases. Understanding self-selection is important for understanding migration processes in general, but its welfare implications only unfold under extreme conditions.

[1] We will summarize this literature in Section 2.

By showing that "*who migrates*" can be as important as "*how many migrate*", this paper also advances the broader literature on the aggregate effects of migration. A series of studies use calibrated general equilibrium models to estimate the impact of migration on GDP per capita. Most of them take the status quo as a benchmark, and estimate the welfare effect of a further reduction in the barriers to international migration (Hamilton & Whalley, 1984; Felbermayr & Kohler, 2007; Klein & Ventura, 2007, 2009; Iranzo & Peri, 2009; Docquier *et al.*, 2012; Kennan, 2013), or take as counterfactual a world without migration (Di Giovanni *et al.*, 2012). Depending on the modeling framework and data, these papers predict significant overall gains from migration. The Mexican example in this paper shows that sizable welfare effects can even arise if the level of migration is kept constant, and the skills of migrants change.

2 The if and why of migrant self-selection: what we know so far

To date, the literature on migrant self-selection has concentrated on two questions: *if* and *why*. Papers focusing on the *if*-question analyze in what characteristics and to what extent migrants differ from the average person in their home country. Papers dealing with the *why*-question try to identify what causes this difference. Both questions are important for understanding migration processes, helping to explain the determinants of migration flows, the outcomes of migrants in the receiving country, and the demographic changes induced by migration.

The theoretical underpinning for studying the causes of self-selection is the Roy model (Roy, 1951), which has been formalized and applied to migration by Borjas (1987). The fundamental driver of migration in this model is the relative returns to skill in sending and receiving countries. A wider income dispersion in the receiving country induces positive selection, because it has the highest benefits for high-skilled migrants. The opposite is true if incomes are more dispersed in the sending country.

While the basic model assumes that a potential migrant knows her income abroad, more recent studies have extended this model. For instance, Bertoli (2010) proves that negative selection becomes more likely once migrants have imperfect information about incomes in the receiving country. Borjas & Bratsberg (1996) show that allowing for return migration reinforces the initial selection pattern.

Additional determinants of self-selection are migration costs, networks, migration policies, and cultural proximity. Migration costs impose a larger hurdle for low-skilled emigrants and lead to a positive selection of migrants (Chiswick, 1999). This effect can be counteracted if migrants have access to migrant networks that lower migration costs and raise the expected income for low-skilled workers (Carrington *et al.*, 1996; Kanbur & Rapoport, 2005; Pedersen *et al.*, 2008; Bertoli & Rapoport, 2013). Selective migration policies can influence selection directly by admitting only certain groups, or indirectly, by making migration more costly for some groups than others. Finally, as shown by Belot & Hatton (2012), closer cultural proximity between sending and receiving country makes it easier for less-skilled workers to migrate, leading to a more negative selection.

One of the most-studied cases in the literature is the self-selection of Mexican emigrants. Drawing on data from the censuses of both countries, Chiquiar & Hanson (2005) conclude that Mexicans are neutrally selected from the Mexican income distribution. Caponi (2010) derives the opposite conclusion, showing that the education distribution of emigrants is U-shaped. Despite arriving at different conclusions, both studies reject the predictions of the Roy model. However, using censuses has the drawback that the same individual cannot be observed in both countries, given that the selection measure can only be based on observable skills. Recently available Mexican panel data, such as the ENET and the MxFLS, allows researchers to observe a person before and after migration, as well as directly computing the skill distributions of migrants and the total population for both observable and unobservable skills. Several studies confirm the Borjas (1987) model, showing that Mexican emigrants are negatively selected on average (Ibarraran & Lubotsky, 2007; Lacuesta, 2010; Fernández-Huertas Moraga, 2011; Ambrosini & Peri, 2012; Kaestner & Malamud, 2013), and that this selection is mainly driven by unobservable characteristics. However, this average masks a significant rural-urban and male-female difference in selection patterns, which is due to wealth constraints, access to migrant networks, and US border enforcement (Orrenius & Zavodny, 2005; McKenzie & Rapoport, 2010; Fernández-Huertas Moraga, 2013).

Besides the US-Mexican case, the forces of the Roy model have been shown to drive migrant selection from many other countries around the world. The evidence ranges from island states in the Pacific (Akee, 2007; McKenzie *et al.*, 2010), middle-income countries in central Europe (de Coulon & Piracha, 2005; Ambrosini *et al.*, 2011; Rosso, 2014) and South America (Bertoli *et al.*, 2010), to the welfare states of Scandinavia (Rooth & Saarela, 2007; Borjas *et al.*, 2013).

Furthermore, differences in the income distribution drive the selection internal migrants. As shown for the US by Borjas *et al.* (1992), people with the highest skills mismatch in a region are most likely to move. In Italy, where returns to skill in the rich North are lower than in the poor South, migrants moving North are negatively selected (Bartolucci *et al.*, 2013). The differences in income distributions also explain the positive selection of rural-urban migrants in China (Xing, 2010), and East-West migrants in

Germany (Brücker & Trübswetter, 2007).

Self-selection was also pervasive in historical migration episodes. Using matched historical censuses from Norway and the US in the late-19th century, Abramitzky *et al.* (2012) find a small positive selection of Norwegian emigrants, although this finding is the sum of a negative selection from urban and a positive selection from rural areas. Similar patterns can be found for returnees from the US. Based on aggregate data from the period 1908-1951, Biavaschi (2012) shows that US out-migrants were initially negatively selected, although the selection became more positive as the US migration policy became more restrictive over time.

In sum, the existing literature provides a detailed picture of the causes — the *if* and *why* — of migrant self-selection, but remains silent on its consequences. In the following sections, we will fill this gap and provide evidence for the aggregate effects of migrant self-selection during two major migration episodes.

3 Migrant self-selection in a model with heterogeneous workers

To determine the welfare impact of migrant self-selection, we rely on a general equilibrium model with heterogeneous workers, based on which we simulate the effect of different self-selection scenarios on the sending and receiving countries. The exercise is a thought experiment, in which we leave the level of migration constant but change the skill composition of migrants, and compare aggregate outcomes under both scenarios. In this research design, the counterfactual is different compared to most studies on the aggregate impact of migration, which change the number of migrants, and consider as counterfactual a world with more or less migration.

Before turning to the analytics of the model, we provide some basic intuition for the simulation exercise. Consider two countries, Mexico and the US. Both are endowed with high-skilled and low-skilled workers, as described in the Edgeworth box in Figure 1. Let A be the endowment of both countries in autarky, that is, before any migration happened. If workers migrate from Mexico to the US, the endowment point moves from A towards the upper right corner within the shaded area. If the endowment after migration lies on the dashed line from A to the upper right corner, migrants are neutrally selected, because the ratio between high- and low-skilled workers is the same for emigrants as the entire Mexican population. Migrants are negatively selected if the new endowment lies North-West of the dashed line, and positively selected if it lies South-West of it. Points B, B', and B", which lie on a 45-degree line, represent migration flows with the same number of migrants, but different selection patterns.

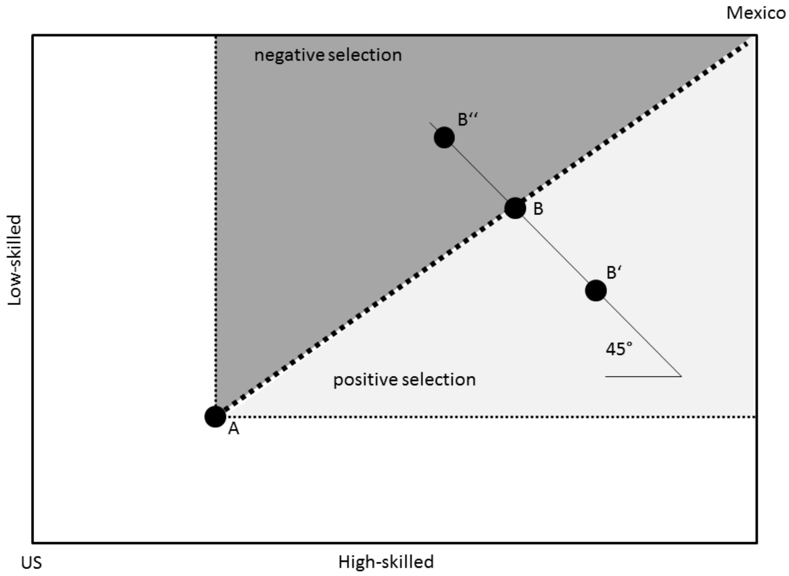

Figure 1: Migration from Mexico to the US.Point A: initial endowments without migration. Points B, B' and B": endowments after migration from Mexico to the US with neutral, positive, and negative migrant selection, respectively.

In the simulation exercise, we compare the economy under the observed migration pattern, for example B', with an economy under neutral selection in point B. This strategy is conceptually different from that applied in other studies, which quantify the difference either between zero migration (point A) and currently observed migration B' (Di Giovanni *et al.*, 2012), or between the current migration B' and a world with more migration, in which the new endowment point lies between B' and the upper right corner (e.g. Docquier *et al.*, 2012; Kennan, 2013). Note that the Edgeworth box implicitly assumes that human capital is perfectly transferable across borders, i.e. that a high-skilled worker in Mexico is also high-skilled in the US. While this assumption is useful to explain the intuition of the research design, we will later relax it, and account for imperfect transferability of human capital as well as differences in skill prices across both countries.

3.1 Basic model

Having laid out the intuition of the research design, we now describe the mechanics of the model. The model is based on heterogeneous workers, allowing us to study both aggregate and distributional effects of self-selection. It closely follows the work of Iranzo & Peri (2009), who use a simplified version of a model developed by Yeaple (2005) to study the aggregate impact of trade and labor market integration in Europe. We will restrict the description of the model to its most important features, and refer the interested reader to Iranzo & Peri (2009) for a full account.

GDP per capita, our main variable of interest, is calculated as the weighted average of real wages.[2] A change in migrant selection affects real wages through two channels: nominal wages and prices. Positive migrant selection makes the workforce in the receiving country more productive compared to neutral selection, leading to a decrease in aggregate prices. At the same time, positive selection increases competition among workers with higher skills, and reduces their nominal wages relative to those of less-skilled workers. As we will show, the productivity effect dominates the competition effect. Consequently, positive selection will increase GDP per capita in the receiving country, while it has the opposite effect in the sending country.

We initially consider each country in autarky, assuming that trade flows do not respond to changes in the skill composition of migrants.[3] A country's total factor productivity (TFP) is denoted by Λ. Each country is populated by a continuum of M workers with skills ranging from the least skilled worker at $Z = 0$ to the most-skilled worker at $Z = 1$. Skills are distributed according to the cumulative density function $G(Z)$. In the sending countries, the initial population M contains all stayers, while M in the receiving countries includes both immigrants and natives. The economy consists of two sectors, X and Y. Sector Y can be understood as the traditional sector, which requires mostly manual-intensive and routine tasks, while sector X is the modern sector, which requires more complex tasks.

Sector Y is perfectly competitive, and produces a homogeneous good with a constant returns to scale technology. Sector X produces N varieties of a differentiated good. Firms can freely enter sector X after paying a fixed cost of F_X units of output. The production technology in sector X exhibits higher returns to skill, g_X, than the technology in sector Y, hence $g_X > g_y$. Workers with a higher skill level Z have a comparative advantage in sector X. As shown by Yeaple (2005), in equilibrium there exists a cutoff skill-level $\bar{Z}$, at which a worker is indifferent between working in sector Y and sector X. Workers with skills higher than $\bar{Z}$ sort into sector X, while workers with skills below $\bar{Z}$ sort into sector Y. $\bar{Z}$ is determined endogenously in equilibrium.

A worker in each sector produces A_Y and A_X units of goods Y and X, respectively, with

[2] We do not model capital, as we are interested in the aggregate long-run effect. Even if capital was included in the model, the long-run outcome would be the same, as capital would fully adjust.

[3] We will relax this assumption in Section 5.3.2.

$$
\begin{aligned}
A_Y(Z) &= \exp(g_Y Z) \\
A_X(Z) &= \exp(g_X Z).
\end{aligned} \tag{1}
$$

Workers are paid their marginal product, such that unit costs are equalized across all skill levels within a sector. Accordingly, the ratio of wage $W(Z)$ and productivity, $A_Y(Z)$ or $A_X(Z)$, is constant within each sector. The worker at the cutoff skill level $\bar{Z}$ is indifferent between working in both sectors, as she receives the same wage in both $W_X(\bar{Z}) = W_Y(\bar{Z})$. In equilibrium, the wage schedule is

$$
W(Z) = \begin{cases} \Lambda \exp(g_Y Z) & 0 \leq Z \leq \bar{Z} \\ \Lambda C_X \exp(g_X Z) & \bar{Z} \leq Z \leq 1 \end{cases}, \tag{2}
$$

with $C_X = \exp(g_Y \bar{Z}) / \exp(g_X \bar{Z}) < C_Y$ being the unit costs in sector X. Good Y is the numeraire, so that $C_Y = P_Y = 1$.

Figure 2 illustrates the wage schedule in equilibrium. The wage schedule is linear in Z, with a kink at $\bar{Z}$ due to the higher returns to skill in sector X. The average nominal wage in equilibrium is a weighted average of all nominal wages,

$$
\bar{W} = \Lambda \left(\int_0^{\bar{Z}} \exp(g_Y Z) dG(Z) + C_X \int_{\bar{Z}}^1 \exp(g_X Z) dG(Z) \right). \tag{3}
$$

To obtain real GDP per capita, $\bar{W}$ has to be divided by the aggregate price index $P = \left[\beta^\theta P_X^{1-\theta} + (1-\beta)^\theta\right]^{\frac{1}{1-\theta}}$, with $P_X = [\int_0^N p(i)^{1-\sigma} di]^{\frac{1}{1-\sigma}}$ being the price index for the differentiated good X.[4]

3.2 Introducing migrant self-selection into the model

We now introduce migrant self-selection into the model and derive predictions for the effect of a change the selection pattern on GDP per capita. Let $G_M(Z)$ be the skill distribution of migrants, and $G_S(Z)$ the skill distribution of the total population in the sending country. We speak of positive selection if migrants have higher skills than the average national of the sending country. Formally, this translates into a first-order stochastic dominance of the migrant skill distribution, $G_M(Z) \leq G_S(Z)$. Migrants are

[4] β is the share of good X in the consumer's utility function, θ and σ are the elasticities of substitution between goods X and Y and between N varieties of X, respectively.

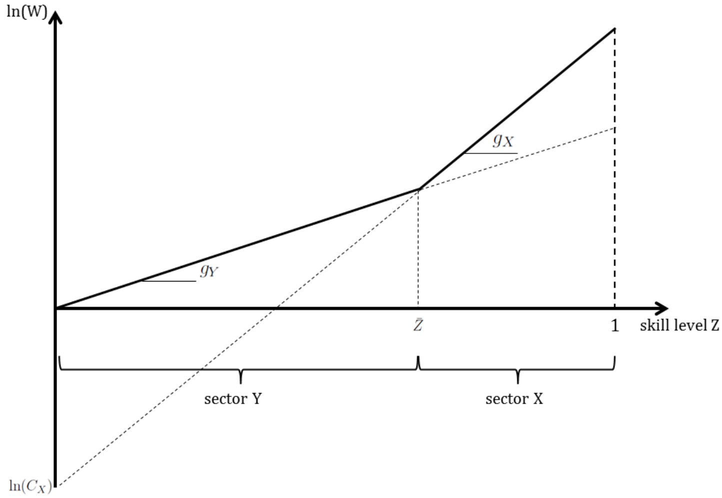

Figure 2: Equilibrium nominal wage schedule.

Notes: See Iranzo & Peri (2009). The equilibrium nominal wage schedule is the upper envelope of the nominal wage schedule in sectors Y and X. Workers self-select into the sector that pays a higher wage. The vertical axis denotes the log nominal wage in terms of the numeraire.

$$
\begin{array}{lll}
\text{positively selected} & \text{if} & G_M(Z) \leq G_S(Z) \quad \forall Z \\
\text{neutrally selected} & \text{if} & G_M(Z) = G_S(Z) \quad \forall Z \\
\text{negatively selected} & \text{if} & G_M(Z) \geq G_S(Z) \quad \forall Z.
\end{array}
$$

As an example, Figure 3 illustrates the effect of negative self-selection on nominal wages in the sending country. The increase in the average skill level of the workforce increases the productivity in sector X, thereby reducing the unit costs of production in sector X. This leads to a downward-shift in nominal wages in the high-skill sector X, and a shift in the cutoff between Y and X to the right. The relative wage decrease in sector X can be interpreted as a competition effect on the labor market. A larger number of high-skilled workers increases competition and reduces nominal wages for higher-skilled workers. At the same time, the sectoral re-allocation from the traditional to the modern sector makes the economy more competitive as a whole, reducing the aggregate price level.

In sum, the effect on real wages depends on the sector. Real wages in sector Y increase due to lower prices, while the effect in sector X can be positive or negative, depending on whether the wage or the price effect dominates. In the receiving country, negative selection has the opposite effect: the total effect on real GDP per capita will be positive, but the magnitude of the effect will depend on the structural parameters of the model.

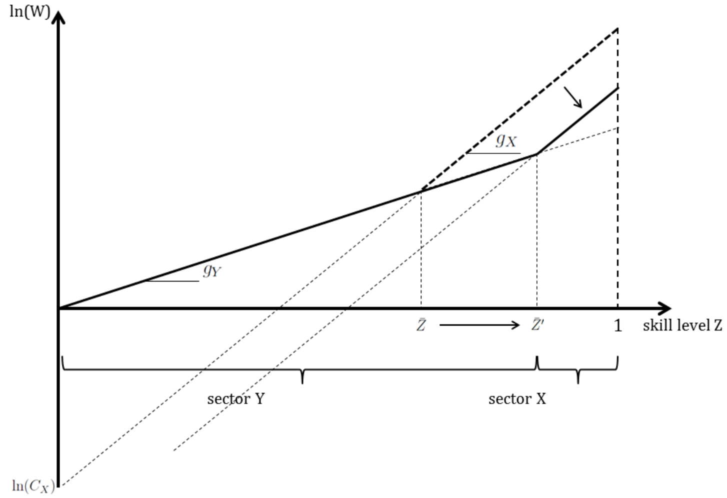

Figure 3: The impact of negative selection in the sending country.

Notes: This figure illustrates the impact of a negative selection on equilibrium nominal wages in the sending countries. If workers become more skilled on average, the cutoff skill level $\bar{Z}$ shifts to the right, leading to lower nominal wages in sector X.

4 Estimating the degree of self-selection

To obtain a measure for the degree of self-selection, we need to estimate the skill distribution of emigrants, $G_M(Z)$, and the total population, $G_S(Z)$. We define an emigrant as a person observed in Norway or Mexico at a given time, who leaves for the US before the following period when his household is surveyed again. A non-migrant is defined as an individual who is observed in one of the sending countries over the entire period. Following the most recent literature (Fernández-Huertas Moraga, 2011; Ambrosini & Peri, 2012; Kaestner & Malamud, 2013), we compute the degree of selection as the difference in pre-migration wages between migrants and the full population. We rely on wages as a proxy for skills for two reasons. First, wages are a reduced-form representation of a worker's human capital, and include observable factors, such as education and experience, as well as unobservable factors, such as motivation and self-confidence. If migrants were positively selected from the sending population, we would expect their higher skill levels to translate into higher wages before migration. By using wages as a skill measure, we can be agnostic about whether selection is driven by observed or unobserved traits. A second advantage of this procedure is that we can directly observe the wages of emigrants, without having to recover their counterfactual wage distribution based on observable characteristics, as in Chiquiar & Hanson (2005) and Biavaschi (2012).

A potential concern with pre-migration wages as a proxy for skills is that wages might decline before migration. If migrants respond to future migration plans by reducing their labor market effort, or by sorting into lower paid occupations, then pre-migration earnings would over- or understate the degree of self-selection. However, Fernández-Huertas Moraga (2011) shows that such concerns are limited for Mexico, and we rely on this assumption in the rest of the paper.

We proceed by briefly explaining our selection measure in each of the source countries and in the US. Full details on the data and variable construction are presented in Appendix A.

4.1 The selection of emigrants from Norway and Mexico

Norway. We first study the migration of Norwegians to the US in the second half of the 19th century. This migration flow is an illustrative example for the mass migration from Europe to the US in the second half of the 19th century, in particular from Scandinavia. While Scandinavian emigration rates were below the European average up to the 1860s, the pattern reversed in later periods, with emigration substantially exceeding European rates (Jensen, 1931). Between 1865 and 1880, the emigration rate from Scandinavia was more than 5 times as large as in the rest of Europe, with Norway driving this pattern. Besides being one of the most important sending countries during the age of mass migration, Norway offers the advantage of having almost completely digitalized censuses. For our analysis, we use the 100% Norwegian Census of 1865, combined with the 1880 US Census, which is the only US Census that has been fully digitalized (Minnesota Population Center, 2008).[5]

We restrict the sample to men between 15 and 40 years old in 1865,[6] and our goal is to attach to each individual an indicator of whether he will have migrated to the US by 1880 - i.e., whether he appears in the US census in 1880. We match the original Norwegian sample in 1865 to a US sample of Norwegian-born males aged 30-55 years, based on an iterative algorithm that has become standard among economic historians (Ferrie, 1996; Abramitzky *et al.*, 2012). In both countries, we first restrict the sample to individuals that can be uniquely identified by first name, last name and age. Names are then standardized to account for orthographic differences. We first match Norwegian men living in Norway in 1865 with Norwegian-born men living in the US in 1880 by name and exact age. If a unique match is found, the observation is considered matched. We then proceed by matching within a one year band around the exact age (additional

[5] Over 95% of Norwegian emigration settled in the US, hence these sources should capture completely the migration flows and their selection pattern during this time period (Jensen, 1931).

[6] We focus on this age group to reduce the risk of not finding individuals in 1880 due to mortality.

details on the matching procedure are available in Appendix A). Migrants are defined as all individuals in the 1865 Norwegian census that we find in the 1880 census, while everybody else is defined as a non-migrant.

Using pre-migration outcomes as a measure of selection complements the evidence given in Abramitzky *et al.* (2012), who compare post-migration outcomes of migrant and non-migrant brothers. The advantage of our strategy is that we do not need to focus on households with multiple siblings; moreover, outcomes for migrants and non-migrants are measured in the same country. Thus, dissimilarities in the occupational distribution of migrants and the total population are not driven by differences in the economic structure of Norway and the US.

Measuring selection in the early censuses poses a further challenge: individual wages are not available. To obtain a wage measure, we assign to each migrant the median income of his occupation. Consequently, selection can only be measured by variation across, but not within occupations. For instance, negative selection should be interpreted as migrants holding lower skilled occupations, although they might be the highest-ability workers within a low-skilled occupation. We use the crosswalk between HISCO occupations and median income provided by Abramitzky *et al.* (2012), who match income levels from Statistics Norway and other sources for 1900 and estimate incomes for more than 200 occupations. The counterfactual distributions are constructed focusing on occupations with an available estimate of average income (about 79.29% of the sample). We standardize the income measure so that its mean is zero, and keep observations within two standard deviations from the mean.[7]

For the simulation exercise, we divide the skill distribution into deciles and calculate the share of migrants and the full population in each decile in 1865. This procedure provides a non-parametric measure of selection that goes beyond differences in mean wages, and captures the impact of self-selection along the entire wage distribution.

To visualize the degree of selection, Figure 4 shows the estimates of the cumulative skill distribution functions of the migrants and the total population, $G_M(Z)$ and $G_S(Z)$. Migrants from Norway were on average mildly positively selected; $G_M(Z)$ stochastically dominates $G_S(Z)$. The Kolmogorov-Smirnov test statistic for equality of both distributions gives a D-statistic of 0.0700, which leads to a rejection of the null hypothesis of equality in the two distributions: migrants' wages are statistically different from those in the total population.

Our analysis confirms the selection patterns found by Abramitzky *et al.* (2012), who show in Table 3 (p. 1847) that migrant selection was on average positive, despite being

[7] This restriction ensures enough dispersion in the distribution that will allow detecting differences between migrants and non-migrants.

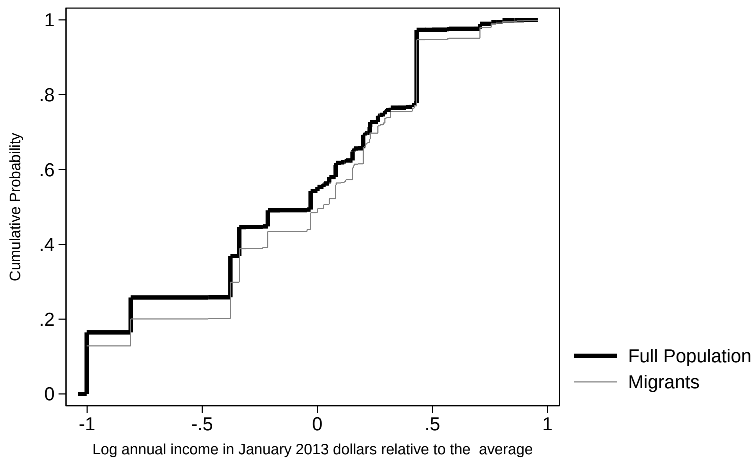

Figure 4: Cumulative Distribution Functions of Migrant and Full Population Skills, Norway 1865

Source: 1865 Norwegian Census.
Notes: Empirical distribution functions of the log of occupation-based median income relative to the annual average income of the full sample. See Appendix A for variable construction.

negative from urban areas. The degree of positive selection is stronger in our data, which might be driven by the fact that we consider an earlier cohort of migrants. Abramitzky *et al.* (2012) consider men aged 3-15 in 1865, who were young enough to be in their childhood household in Norway and were found in the 1900 Censuses, while we focus on men aged 15-40 in 1865, who were young enough to be in the labor force in both 1865 and 1880. Falling transport costs and the greater importance of migrant networks are possible reasons why the positive selection became less pronounced over this period (Hatton & Williamson, 1998), and why we find a stronger degree of selection than Abramitzky *et al.* (2012). To ensure that our results are fully consistent with those in Abramitzky *et al.* (2012), we link men aged 15-25 in the 1875 Norwegian census (who would be in Abramitzky et al.'s sample) to Norwegian-born migrants in the 1880 US Census. Additionally, the shorter time span between the census rounds reduces the role of selective mortality in causing non-matches. The results from this exercise are comparable to those in Abramitzky *et al.* (2012), with positive selection being slightly smaller and opposite selection patterns between rural and urban areas. However, the overall conclusions of this paper are not affected if we use this other sample.

Mexico Mexicans accounted for the majority of migrants in the most recent wave of mass migration to the US, with their degree of selection having been intensively debated in the literature (Chiquiar & Hanson, 2005; Ibarraran & Lubotsky, 2007; Lacuesta, 2010;

Fernández-Huertas Moraga, 2011; Ambrosini & Peri, 2012; Kaestner & Malamud, 2013).

To estimate the degree of selection of Mexican emigrants, we closely follow the work of Fernández-Huertas Moraga (2011). We use the Encuesta Nacional de Empleo Trimestral (ENET) from the second quarter of 2000 until the third quarter of 2004. This nationally representative survey follows household for five quarters, and includes information on household members who left for the US. The information on emigrants is given by the remaining household members in Mexico, which means that we do not observe migrants whose entire household migrated.[8] As with Norwegians, we define Mexican emigrants as individuals who are present in Mexico at the time of the survey and who are reported to have migrated to the US in the following quarter. Non-migrants are individuals who are observed in Mexico throughout the sample period.

The final dataset comprises all survey rounds from 2000 to 2004, with an identifier for all people who migrate in the quarter following the survey date. We restrict the sample to men between 25 to 65 years with non-missing wage information, working between 20 and 84 hours per week. As before, we focus on individuals with wages within two standard deviations from the mean in constructing the wage distributions.

Applying the same procedure as for the Norwegian sample, we estimate the skill distribution for migrants and the full population based on pre-migration hourly wages. Figure 5 shows $G_M(Z)$ and $G_S(Z)$. The cumulative skill distribution of migrants lies above that of the full population, indicating that migrants are negatively selected. The Kolmogorov-Smirnov test statistic for equality in the migrant and counterfactual distribution gives a D-statistic of 0.0882, rejecting the null hypothesis that migrants are drawn from the same distribution of the full population.

4.2 Migrant selection and the skill distribution in the US

To quantify the aggregate impact of migrant selection in the US, we require a counterfactual skill distribution, namely one that would occur if immigrants were neutrally selected from their home country population. Compared to the sending countries, obtaining a counterfactual for the US is challenging, because we do not observe the skills of neutrally selected immigrants. In the sending countries, the same skill prices apply to both

[8] To avoid this problem, Kaestner & Malamud (2013) use the Mexican Family Life Survey (MxFLS), where migrants are not only identified by the households left behind, but are followed across borders. These authors confirm the negative selection of male migrants. For our purposes, the main disadvantage of the MxFLS is the much reduced sample size: the pre-migration wage distribution would be based on a sample of about 200 migrants. However, complementary analyses based on this second dataset yield conclusions that are in line with those presented in this paper.

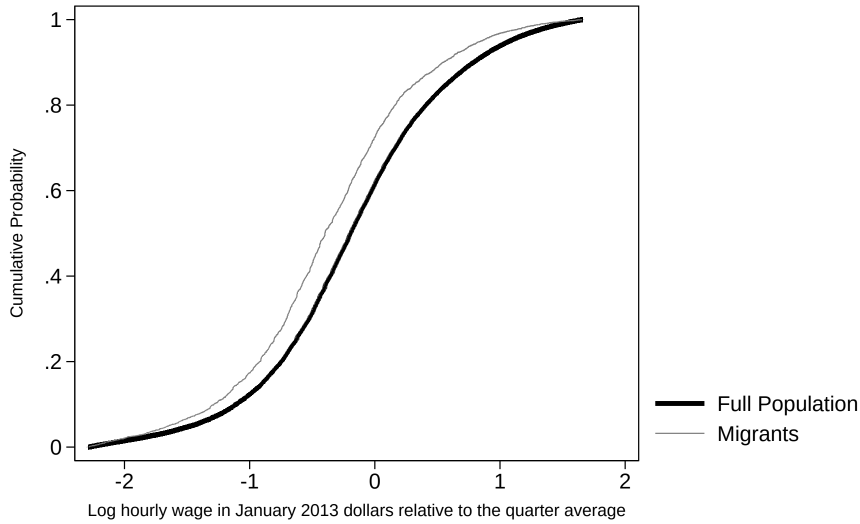

Figure 5: Cumulative Distribution Functions of Migrant and Full Population Skills, Mexico 2000-2004

Source: ENET.
Notes: Empirical distribution functions of the log-hourly wages relative to the average wages of the full sample in a given quarter. See Appendix A for variable construction.

emigrants and the full population, such that the skill distribution of the full population can be used as a counterfactual. In the US, we cannot simply apply the counterfactual skill distribution we found for Norway and Mexico, because skills are rewarded differently across countries and human capital acquired in the home country cannot be easily transferred to the US. To account for these issues, we construct the skill distribution of neutrally selected migrants in the US by applying a re-weighting procedure similar to DiNardo *et al.* (1996) and Chiquiar & Hanson (2005).

Before turning to the counterfactual, we first present stylized facts about the baseline skill distributions of natives and current migrants in the US. For 1880, we use the full US Census, restricting the sample to men between 15 and 40 years old. The income variable represents the median income by occupation in 1950, which allows us to separately rank individuals of the migrant population and the full population using a consistent definition for more than 200 occupations. All income variables are inflated to 2013 US dollars. For the US in 2000, we use the 5% sample of the US census, available from IPUMS. We restrict the sample to males between 25 and 65 years old and currently working between 20 and 84 hours per week, and we construct hourly wages as the ratio of annual income and usual hours worked per year.

The first two panels of Figure 6 and Figure 7 shows the cumulative distribution function (cdf) of US natives, as well as Norwegian and Mexican immigrants in the US. Both migrant groups have lower skills on average than US natives, although Norwegian migrants at the bottom of the skill distribution outperform US natives. The skill difference between immigrants and natives is considerably larger for Mexicans than Norwegians. The Kolmogorov-Smirnov D-statistic is 0.0695 for the 1880 and 0.3593 for the 2000 sample, indicating that migrants and natives significantly differ in their skills. The cdfs for Norwegian and Mexican immigrants in Figure 6 will serve as baseline scenarios in the simulation exercise.

We now turn to the counterfactual skill distribution in the US. For the sake of clarity, we will discuss here the example of Mexican immigrants in 2000, but obviously the same arguments apply to Norwegian immigrants in 1880. The counterfactual skill distribution that we would like to recover is $g^{US}_{neutral}(w|Z)$, the distribution of wages conditional on skills Z that would be observed in the US if Mexican immigrants were neutrally selected from the full Mexican population. It can be expressed as

$$g^{US}_{neutral}(w|Z) = \int f^{US}(w|Z)h(Z|US, neutral)dZ, \tag{4}$$

where $f^{US}(w|Z)$ is the density of wages w in the US conditional on skills Z, and $h(Z|US, neutral)$ represents the skill distribution of neutrally selected migrants in the US. The challenge in estimating (4) is that $h(Z|US, neutral)$ is unobserved.[9]

If skill prices were the same in the US and Mexico and skills were fully transferable, the counterfactual distribution in the US would be equal to the skill distribution of the full population in Mexico, as shown by the solid lines in Figure 4 and 5. In this case, the counterfactual skill distribution would be written as

$$g^{US}_{neutral}(w|Z) = \int f^{US}(w|Z)h(Z|Mex, neutral)dZ. \tag{5}$$

However, given the economic and institutional differences between the US and Mexico, Equation (5) would be a naïve estimator for the counterfactual skill distribution, vastly over-estimating the impact of migrant selection. For example, it would assume that a Mexican who is in the 9th decile of the wage distribution in Mexico will be in the 9th decile of the US wage distribution. This is unrealistic because human capital acquired in Mexico cannot be fully transferred across borders, which is why migrants often work in jobs for which they are over-educated (Piracha & Vadean, 2013). Moreover, the same

9 We choose this notation to make our approach comparable to DiNardo *et al.* (1996) and Chiquiar & Hanson (2005). Strictly speaking, we would not need the wage density $f^{US}(w|Z)$ in the equation, as our initial skill measure equals wages. Without loss of generality, we could assume that $f^{US}(w|Z) = 1$.

skills might be rewarded differently in the US and Mexico, while migrants and natives with the same observable characteristics are not necessarily perfect substitutes in the US labor market (Borjas *et al.*, 2008; Ottaviano & Peri, 2012). Consequently, the skill distribution of emigrants in Mexico can have a completely different shape compared to that of the same migrants in the US. For the same reason, we would expect the counterfactual skill distribution in the US to have a different shape than that in Mexico. Given the negative selection of Mexican emigrants, we would expect the locus of the counterfactual skill distribution to the right of the currently observed immigrant skill distribution in the US, although the difference between both should be small.

To make the skills of neutrally selected immigrants comparable to those of the currently observed migrants in the US, we apply a weighting strategy similar to Chiquiar & Hanson (2005). Chiquiar & Hanson construct a counterfactual that allows them to compare the skills of Mexican migrants currently in the US with those of non-migrants in Mexico, thus determining how these migrants would fare if they were working in Mexico. Our procedure works the other way round; we take the skill distribution of neutrally selected emigrants in Mexico and determine how these migrants would fare in the US.

To this end, we choose weights that re-adjust the observed skill distribution of migrants currently in the US to account for differences in skills driven by migrant self-selection. $g^{US}_{neutral}(w|Z)$ in Equation (4) can be re-written as a weighted average of the observed skill distribution for negatively selected migrants in the US

$$g^{US}_{neutral}(w|Z) = \int \theta f^{US}(w|Z) h(Z|US, neg) dZ. \tag{6}$$

The weighting factor θ takes the form[10]

$$\theta = \frac{h(Z|US, neutral)}{h(Z|US, neg)} = \frac{\Pr(US, neutral|Z)}{\Pr(US, neg|Z)}. \tag{7}$$

The numerator of θ gives the proportion of neutrally selected migrants at every skill level Z, while the denominator captures the proportion of negatively selected migrants with skills Z. While $h(Z|US, neutral)$ obviously remains unobservable, it is possible to obtain an estimate for the ratio of conditional probabilities in Equation (7), and thus for θ.

In practice, we use the ratio of conditional densities of neutrally and negatively selected migrants in Mexico as a non-parametric estimate for θ, which we replace in Equation (7) with a new weight

$$\hat{\theta} = \theta_{Mexico} = \frac{Pr(Mex, neutral|Z)}{Pr(Mex, neg|Z)}. \tag{8}$$

[10] See Equations (8)-(16) in Chiquiar & Hanson (2005) for the derivation of the weighting factor.

For each decile of the Mexican skill distribution, we have obtained an estimate of the density of negatively selected emigrants, $Pr(Mex, neg|Z)$ and the share of the full population, $Pr(Mex, neutral|Z)$, such that we can compute θ_{Mexico} for each decile of Z. Applying the weights from Equation (8), we obtain the counterfactual skill distributions shown in Figure 7. For comparison, Figure 7 also displays the unweighted counterfactual distribution, which replaces $h(Z|US, neutral)$ with $h(Z|Mex, neutral)$. Imposing neutral selection brings relatively more low-skilled Norwegian migrants and relatively more high-skilled Mexican migrants to the US compared to the observed migrant skill distribution. Re-weighting reduces these differences, especially in Mexico.

This re-weighting strategy is based on the assumption of rank insensitivity across countries (Dustmann *et al.*, 2013), that is, the assumption that the relative ranking of migrants in the home country is preserved in the US. If Mexican A has higher skills in the Mexican labor market than Mexican B, $Z_A^{Mex} > Z_B^{Mex}$, rank insensitivity assumes that A also has higher skills in the US labor market, $Z_A^{US} > Z_B^{US}$. Rank insensitivity assures that these two inequalities hold even if imperfect human capital transferability and differences in skill prices compress the immigrant skill distribution in the US and lead to a smaller skill gap between both individuals in the US, $Z_A^{US} - Z_B^{US} < Z_A^{Mex} - Z_B^{Mex}$. Through this assumption, we can exploit the relative difference between migrants and the full population in Mexico bearing informational content that can be used to project the relative share of migrants over the full population onto the US skill distribution.

To further clarify the mechanics of our reweighing procedure, consider the following example. Suppose that 20% of all emigrants, but only 10% of the full population, is in the first decile of the Mexican skill distribution — as measured by pre-migration earnings. This proportion tells us that negatively selected emigrants are twice as likely to be in the lowest decile of the skill distribution than neutrally selected emigrants. In this case, $\theta_{Mexico} = 10\%/20\% = 0.5$, which we take as the weight for the first decile in the US skill distribution. Suppose that the share of negatively selected migrants in the lowest decile of the US skill distribution is 30%, with skill downgrading and different skill prices pushing more migrants to the bottom of the skill distribution compared to what is observed in Mexico. Applying the weights θ_{Mexico} as in Equation (6), we now have $30\% * 0.5 = 15\%$ of neutrally selected immigrants in the lowest decile in the US.

This procedure deviates from Chiquiar & Hanson (2005), in both the measurement of skills, and the estimation of the weights θ. Chiquiar & Hanson focus on observable skills such as age, education, marital status, and residence in a metropolitan area, which is available in the censuses of Mexico and the US.[11] Based on these skills, they use a logit

[11] However, Ibarraran & Lubotsky (2007) and Fernández-Huertas Moraga (2011) show that the combination of these two data sources strongly distorts the pattern of selection due to differences in

model to predict whether a Mexican is in the US, which provides a parametric estimate for the weights θ.

In comparison, our data allow us to rely on a more comprehensive skill measure, given that we can fully exploit the information on unobservable characteristics contained in pre-migration earnings. An additional advantage of the data is that we observe migrants and non-migrants in the same labor market, which avoids the complications stemming from the combination of different data sources. Finally, our estimator for θ is fully non-parametric, such that the counterfactual skill distribution will be based on both observable and unobservable factors that determine wages in the US.

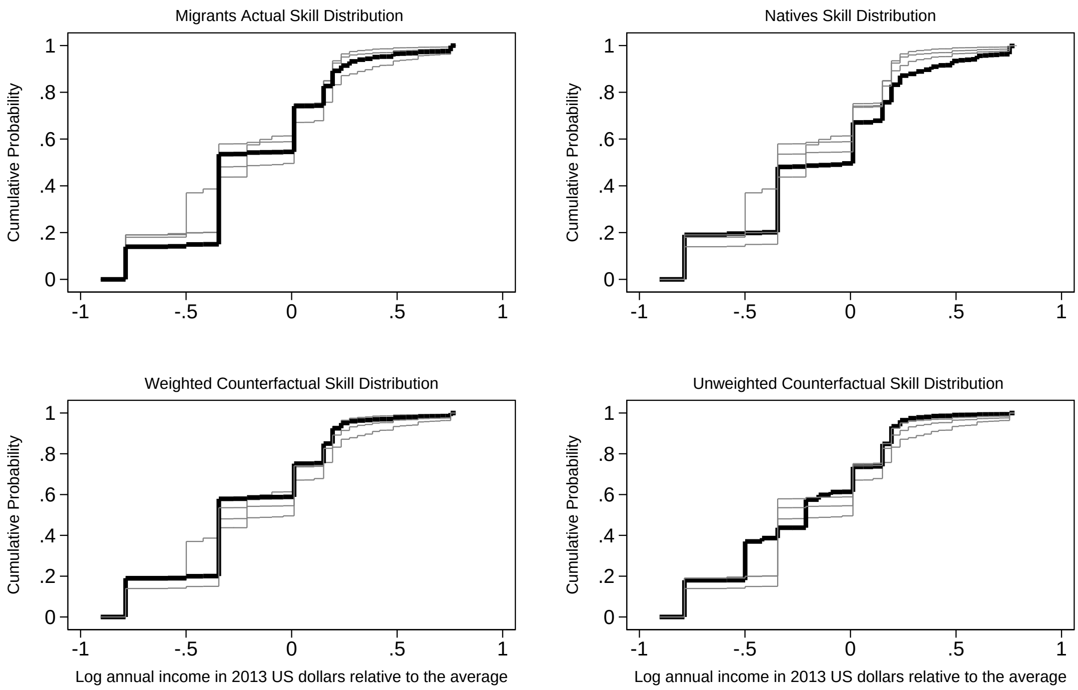

Figure 6: Actual and Counterfactual Cumulative Distribution Functions of Natives, and Norwegian Migrant Skills, US 1880.

Source: 1880 US Census.
Notes: Each figure displays four skill distributions: one for natives, and three for immigrants. The "actual" distribution is the observed skill distribution of immigrants in the US, whereas the counterfactual distributions are the skill distributions of neutrally selected immigrants in the US. The "unweighted" distribution simply uses the counterfactual distribution from Norway, while the "weighted" distribution applies the weighting procedure specified in this section.
All figures present the empirical distribution functions of the log of occupation-based mean income relative to the annual average income of the full sample.
See Appendix A for variable construction.

sampling procedures and variable definitions. Therefore, the combination of the two censuses might pose serious challenges to the estimation of equation (4). Additionally, by only conditioning on observable characteristics, the recovered distribution would not account for self-selection on unobservables.

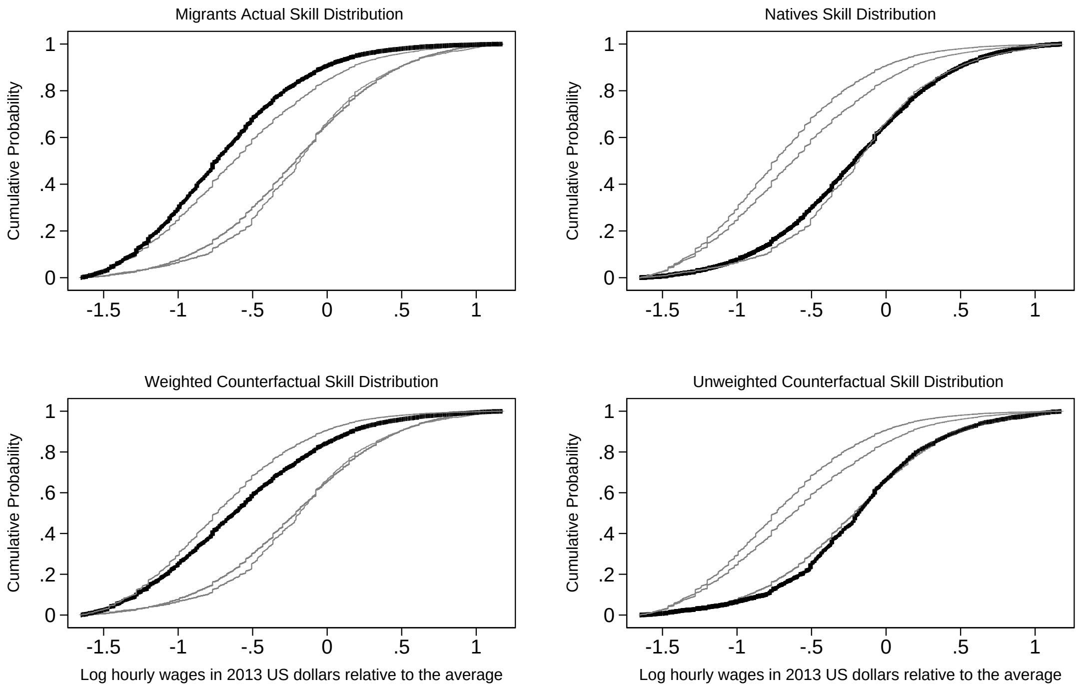

Figure 7: Actual and Counterfactual Cumulative Distribution Functions of Natives, and Mexican Migrant Skills, US 2000.

Source: 2000 US Census.
Notes: Each figure displays a skill distribution: one for natives, and three for immigrants. The migrants "actual" distribution is the observed skill distribution immigrants in the US. The counterfactual distributions are the skill distributions of neutrally selected immigrants in the US. The "unweighted" distribution simply uses the counterfactual distribution from the sending countries, while the "weighted" distribution applies the weighting procedure specified in this section.
All figures present the empirical distribution functions of the log-hourly wages relative to the quarter average.
See Appendix A for variable construction.

5 The aggregate impact of migrant self-selection

Equipped with the estimates for the degree of self-selection for both sending countries, we now turn to the simulation exercise. In a thought experiment, we compare aggregate outcomes under the observed selected migration with a scenario in which migration occurs at the same level, but migrants are neutrally selected. As a first step, we calibrate the model outlined in Section 3 on the economies of Norway and the US in 1880, as well as Mexico and the US in the early 2000s. We then feed in the estimated skill distributions of migrants and the full population from Section 4, and calculate the difference in aggregate outcomes under different migration scenarios.

5.1 Calibration

We calibrate the model such that it replicates the baseline features of the economies of the sending and receiving countries. Most parameters are taken from the literature or calculated from available data sources. For parameters that cannot be calculated, we pick values such that key moments generated by the model match the real world data. The calibration is summarized in Table 1.

Table 1: Parameters for calibration

	Parameter	Norway 1865	USA 1880	Mexico 2000s	USA 2000s
External parameters					
Returns to skill, sector Y	g_Y	1	1	1	1
Population: stayers, natives		1,702	43,391	101,826	244,000
Emigrants to US		183		10,017	
Immigrant stock, all countries			1,414		26,588
TFP	Λ	0.3195	1	0.286	1
Internal parameters					
Fixed cost in sector X	F_X	5,980	21,717	40,677	19.5
Returns to skill, sector X	g_X	1.02	1.48	2.83	1.42
Preference share of X	β	0.57	0.43	0.46	0.52
Elasticity of substitution X and Y	θ	1.66	1.61	1.32	1.57
Elasticity of substitution, varieties of X	σ	4.07	4.34	3.82	4.21

Note: Population and migrant numbers in 1000s. External parameters are computed from external data sources or taken from other studies. Internal parameters are estimated within the model, using an SMM procedure.

The population is measured as the number of non-migrants in the sending countries and natives in the receiving countries. The migrant numbers are taken from the censuses of the receiving countries. The sources are the US census in 1880 and the 2000 US census for the number of natives and immigrants, the Norwegian census in 1865 for non-migrants in Norway, and the OECD population statistics in 2002 for non-migrants in Mexico. TFP in the US is normalized to one. The TFP level for Norway is based on Williamson (1995). For the difference in TFP between Mexico and the US in the 2000s, we use the differences in labor productivity levels provided by the OECD. g_Y, the returns to skill in sector Y, are normalized to 1.

The sectoral classification is based on the ISCO (HISCO in our historical samples, van Leeuwen *et al.*, 2002) standardized occupational classifications.[12] We choose the fixed costs F_X, the returns to skill g_X, and the preference parameters β, θ and σ, such that the model outcomes match the observed shares in the two sectors, as well as quintiles of

[12] Skilled occupations are legislators, senior officials and managers, professionals, technicians and associate professionals, clerks, service workers and shop and market sales workers. Unskilled occupations are skilled agricultural and fishery workers, craft and related trades workers, plant and machine operators and assemblers, elementary occupations .

the observed nominal wage distribution. Appendix B provides details on the numerical procedure used to compute these internal parameters.

5.2 Simulation results

Based on the calibration shown in Table 1, we now simulate the changes in the migrant skill distribution, and calculate its effects on sending and receiving countries. Before turning to the results, let us recall the theoretical predictions. For the receiving countries, a more positive selection of migrants leads to a more highly skilled workforce, resulting in a lower price level and a higher level of GDP per capita. In light of these predictions, we would expect a positive impact on GDP per capita in the US in 1880 and Mexico in the 2000s, and a negative impact on Norway in 1865 and the US in the 2000s.

While the theory predicts the sign of the aggregate effect, its size depends on the model parameters. The same holds true for the share of workers in both sectors. While an increase in the skill-level shifts the cutoff between the low-skilled sector Y and the high-skilled sector X to the right, this might not necessarily translate into a higher share of workers in sector Y, given that the skill distribution shifts at the same time. Therefore, if there is a strong increase in the share of high-skilled workers, it is possible to have a higher cutoff and a higher share of workers in sector X.

The magnitude of the aggregate effects will depend on both the share of migrants, and the degree of selection. If the share of migrants is low compared to a country's population, the effects will be small regardless of the skill composition of the migrant flows. Given that people migrated from a smaller to a larger country in both cases, the relative population changes are naturally larger in the sending countries. As we can see in Table 1, the emigration of 183,000 Norwegians reduced the Norwegian population by 9%, while it only increased the US population by 0.4%. Likewise, the emigration of 10 million Mexicans to the US reduced the Mexican population by 9%. The corresponding increase in the US population was smaller, but still amounted to 4.1%. Given the differences in population changes, we can expect the effects in Norway and Mexico to be larger than in the US.

Another determinant of the effect size will be the degree of migrant selection. The effects will be small if migrants have almost the same characteristics as all nationals from the country of origin, that is, if baseline and counterfactual skill distributions are almost identical. As shown in figures 4 and 5, the degree of selection was a lot smaller in Norway in 1865 than Mexico in the 2000s. Accordingly, we would expect larger effects for Mexico, even if all other parameters were the same as in Norway in 1865.

The gains and losses from migrant self-selection, computed as the relative difference between baseline and counterfactual, are presented in Table 2. Every entry of this table is the change in a variable that was caused by migrant self-selection. For example, the change in Norwegian GDP per capita in Column (1) means that Norwegian GDP per capita in 1865 was 0.26% lower due to a positive selection of emigrants than it would be under a neutral selection of emigrants.

Table 2: The gains and losses from migrant selection

	Sending countries		Receiving countries	
	Norway 1880	Mexico 2000s	USA 1880	USA 2000s
	(1)	(2)	(3)	(4)
Changes in %				
Real GDP p.c.	**-0.26**	**1.11**	**0.02**	**-0.28**
Price index	-0.09	-0.64	-0.00	0.13
Changes in percentage points				
Gini	0.02	0.00	-0.00	-0.04
Employment Y	-0.25	-0.15	-0.00	0.08

Notes: The table displays the results of the simulation exercise. Each number is the percentage or percentage point difference between selective migration and neutrally selected migration.

The aggregate effects for the sending countries are shown in the first two columns. Besides reducing GDP per capita, Norwegian emigration reducing the price level by 0.09%. Overall, the effect of migrant selection on GDP per capita in Norway is small, owing to the low degree of selectivity.

By contrast, the effects in Mexico are considerably larger. While the negative selection of Mexican emigrants contributed little to wage inequality and the sectoral distribution, it had a significant impact on aggregate prices and real GDP per capita. Because the 10 million Mexican emigrants were less skilled than the average Mexican, the country is left with a more productive workforce, and Mexican GDP per capita is 1.11% higher. The effect is much larger in Mexico than it was in Norway, mainly because the degree of selection was larger in Mexico.

In the US, the aggregate effects are close to zero in both periods, as shown in Columns (3) and (4) of Table 2. The inflow of 183,000 Norwegians in the 1880s only translated into an incremental increase of the US population, and thus the composition of these flows did not matter at the aggregate level. Moreover, although immigration from Mexico 120 years later occurred on a much larger scale, the negative selection of Mexicans only had a small effect, reducing US GDP per capita by 0.28%. This result might be surprising given that Mexican immigration increased the US population by over 4% and given the significant negative selection. However, it can be explained by imperfect transferability of human capital, which leads to a narrow skill dispersion of Mexican immigrants in the

US. Despite migrants being drawn from all parts of the skill distribution in Mexico, after emigration they do not end up in the same decile of the US skill distribution. On the contrary, Mexicans are heavily concentrated at the lower end of the US skill distribution, as shown in the first panel of Figure 7. Therefore, even if the selection of migrants changes significantly with respect to the skill distribution in Mexico, it only marginally changes the skill distribution of Mexicans in the US.

Taken together, the results suggest that migrant self-selection can — but does not necessarily have to — matter on the aggregate. While many studies on the causes of self-selection are motivated by the potential welfare consequences in the receiving countries, our results show that the effects in the US are close to zero. However, the effect is economically significant in the sending countries, especially in Mexico. The outflow from Mexico was large, and Mexican migrants differ considerably from the full population in terms of their characteristics, which translates into a substantial effect on GDP per capita. In the next section, we will check the robustness of these results to changes in the model assumptions, and assess the size of the effects.

5.3 Extensions and further discussion

We now provide further insights on the magnitude and sensitivity of the effects found in the previous section. We first put the large effect in Mexico into perspective, by comparing it with the pure scale effect, that is, the difference between zero migration and the current level of migration. Furthermore, we extend the model with international trade, and show how the simulation results change when migration alters both the skill distribution and the output mix of the economy. We also derive an upper bound for the aggregate effect in the US, by assessing the impact on the US economy if all immigrants were selected like Norwegians or Mexicans. Finally, we discuss the sensitivity of the results with respect to human capital externalities, as well as the extent to which the change in GDP per capita can be interpreted as a welfare effect.

5.3.1 Who vs. how many? How important is the aggregate effect of selection?

The simulation results show that migrant self-selection can have a significant impact on income per capita, provided that the size of the migration flow and the degree of selection are sufficiently large. Among the four cases we consider, the positive effect in Mexico stands out. Mexico is one of the largest migrant sending countries in the world, which is why we want to take a closer look at the consequences of its migrant selection. Because Mexicans are largely negatively selected, GDP per-capita is about 1%

Table 3: Simulation results: extensions

	Scale effect		Trade response		All migrants selected	
	USA 2000s	Mexico 2000s	USA 2000s	Mexico 2000s	USA 1880	USA 2000s
	(1)	(2)	(3)	(4)	(5)	(6)
Changes in %						
Real GDP p.c.	**-0.99**	**1.01**	**-0.37**	**1.85**	**0.14**	**-0.94**
Price index	0.47	-0.58	0.16	0.16	-0.03	0.41
Changes in percentage points						
Gini	-0.18	0.00	-0.07	0.00	-0.02	-0.10
Employment Y	0.35	-0.14	0.32	-1.70	-0.00	0.19

Note: This table presents extensions to the basic simulations. Columns (1) and (2): scale effect of migration, i.e. the difference between zero migration and the currently observed level of migration. (3) and (4): effect of selection with costless trade. (5) and (6): aggregate effect in the US if all migrants were selected like Norwegians/Mexicans.

higher than it would be if migrants were neutrally selected. However, since we had no strong prior about the size of this effect, we would like to assess whether this increase is economically significant.

To assess the size of the selection effect, we compare it to an objective benchmark that has been widely studied in the literature — the aggregate impact of migration *per se.* While our selection effect purely measures the impact of *"who migrates"* — changing the skill composition while leaving the number of migrants constant — most of the literature estimates the impact of *"how many migrate"* — leaving the skill composition constant while changing the scale of migration from zero to 10 million. In the Edgeworth box in Figure 1, the scale effect is the difference between A and B', while the selection effect is the difference between B and B'.

Columns (1) and (2) of Table 3 display the simulation results for the pure scale effect, which compares the Mexican economy under the currently observed level of migration to a counterfactual economy without migration. The results show that the selection effect greatly matters in Mexico, while it is modest in the US. In Mexico, the selection effect of +1.11% is even larger than the scale effect of +1.01%. To the same extent that it matters if 10 million Mexicans leave the country, it also matters that these 10 million have different skills compared to the full Mexican population. By contrast, the scale effect in the US is almost four times larger than the selection effect. It matters almost four times as much that 10 million Mexicans are in the US than it does that these are less-skilled than the average Mexican. Yet, if the US were to entirely lose 10 million Mexicans, this would have a significant negative effect on US GDP per capita.

5.3.2 Accounting for trade responses

The benchmark simulations in Section 5.2 were carried out for each country in autarky. A change in migrant selection results in a pure surplus effect; the workforce becomes more or less productive, which changes the level of GDP per capita. Nonetheless, a change in the skill composition of migrants also changes the relative skill endowment of each economy. For example, the negative selection of Mexicans means that Mexico becomes relatively more skilled and the US relatively less. In autarky, the economy adjusts to changes in the skill composition of the workforce with a shift in the sectoral distribution of skills, yielding a positive effect in Mexico, given that it is left with a more productive workforce, and a negative effect in the US.

If both countries trade with each other, changes in the skill endowment have an additional effect on the sectoral distribution of skill types, as it might induce a change in the specialization pattern. Suppose Mexico is initially specialized in agricultural production, having a relatively larger Y-sector than the US. Accordingly, the negative selection of emigrants has two effects: a positive effect because Mexican workers become more productive on average, and a change in the trade patterns, because the higher-skilled workforce makes the production in sector X more profitable, shifting part of the workforce from sector Y to sector X.

To assess the robustness of our results when we allow for trade, we compare the aggregate effect under autarky to the effect in an economy with costless trade. While these are polar cases, they provide bounds to the actual effect. In Appendix E, we explain in detail how we incorporate trade into the baseline model. For calibration, we use the same population numbers, immigrant numbers, fixed costs, and returns to skill as in the autarky case. For better tractablility, we now assume that the preference parameters are the same across countries, and choose similar values to those obtained in Section 5.1: $\beta = 0.5$, $\theta = 1.45$, and $\sigma = 4$.

As Columns (3) and (4) in Table 3 show, under autarky we actually under-estimate the impact of self-selection. Under costless trade, the effects on GDP per capita are larger compared to the benchmark. The effect in Mexico is $+1.85\%$ instead of $+1.11\%$ in autarky, while the effect in the US is now -0.37% as opposed to -0.28% in the benchmark. The results with trade confirm that the selection of migrants changes the specialization pattern of both economies. In Mexico, the employment in sector Y decreases by 1.7 percentage points, while it increases by 0.44 percentage points in the US, leading to an additional impact of selection on GDP.

5.3.3 What if all immigrants were selected like Norwegians/Mexicans?

So far, our results demonstrate that the selection of Norwegians in 1865 and Mexicans in the 2000s has almost zero impact on the US economy. However, both groups are only two among many immigrant groups in the US. While the selection of one group might not matter for the US as a whole, it would potentially matter if all immigrants in the US were selected according to a given pattern. To quantify this effect, we conduct the following exercise in the US in 1880 and the 2000s: suppose that all immigrants in the US were selected like Norwegians or Mexicans, what would be the aggregate effect on US GDP per capita?

Even if we impose the same selection pattern on all US immigrants, the effects remain small. As shown in Columns (5) and (6) of Table 3, the impact on real income per capita is larger by a factor 7 in 1880 and a factor 3 in 2000, but still only amounts to a 0.14% increase in 1880 and a 0.94% decrease in 2000.

These findings suggest that self-selection mostly matters for the sending countries. In the US, given the concentration of immigrants in the lower end of the skill distribution, and the narrow skill dispersion of immigrants, even a significant change in their selection pattern does not lead to large effects at the aggregate level.

5.3.4 Human capital externalities

The model outlined in Section 3 calculates GDP per capita as the weighted average over the real wages of the entire population, multiplied by TFP. Our model abstracts from human capital externalities. TFP is assumed to be constant, such that a change in the selection pattern leads to a proportional change in GDP per capita, regardless of the initial level of TFP. If we followed Lucas (1988) and included human capital externalities into the model, the impact of skills on GDP per capita would be non-linear. Equation 3 would then become

$$\bar{W} = \Lambda(\tilde{Z}) \left(\int_0^{\bar{Z}} \exp(g_Y Z) dG(Z) + C_X \int_{\bar{Z}}^1 \exp(g_X Z) dG(Z) \right), \tag{9}$$

where $\tilde{Z}$ is the average level of human capital in the economy. Suppose $\Lambda(\tilde{Z})$ was monotonically increasing in $\tilde{Z}$, then an increase in the level of skills would have two effects on GDP per capita. As before, it would increase real wages because workers are more productive on average. As a second-order effect, the production technology itself would now become more efficient, which would lead to an additional increase in GDP per capita.

While including human capital externalities might be appealing for studying long-run economic development, we believe that leaving TFP constant is a sufficient approximation for the human capital changes induced by migrant self-selection. Consider Mexico, which had the highest share of migrants among the four examples. By moving from negative to neutral selection, we replace 10 million Mexicans with low skills with 10 million Mexicans with slightly higher skills, while the skills of 91% of all Mexicans remain constant. Moreover, even if the human capital externality mattered, as long as TFP increases in the average level of human capital, the estimates in Table 2 would reflect a lower bound to the true effects.

5.3.5 GDP per capita changes, but what about welfare?

Our results show that migrant self-selection affects the level of GDP per capita. However, one might be concerned that an increase in GDP per capita may not translate into an increase in welfare for the full population. According to this line of argumentation, the impact on GDP per capita could be interpreted as a mere statistical effect. This would be the case if all the gains or losses would go to migrants, while all non-migrants would be unaffected.

While being a serious concern, a glance at the aggregate effects in Table 2 shows that it is unfounded. A change in human capital mainly affects GDP per capita through aggregate prices. If workers are on average more highly-skilled, they can produce more output per capita, which ceteris paribus reduces the aggregate price level. In Mexico, for example, negative selection reduces the price index by 0.64%, which is the same for all non-migrants. Consequently, a change in the skills of 9% of the Mexican population benefits the remaining 91%, because their goods become cheaper and they can afford more consumption. Moreover, even if the surplus created by self-selection only accrues to migrants, non-migrants could be made better off through redistribution. In sum, the predicted change in GDP per capita is not merely a statistical effect, but rather reflects a change in the welfare of the entire population.

6 Conclusion

Migrant self-selection is a central theme in the economics of migration. A large amount of literature focuses on the causes of self-selection, investigating if and why migrants differ from the average person in their home country. In this paper, we turn to the consequences of migrant selection, and ask whether it actually matters for the sending and receiving countries. To quantify the effect of migrant selection on GDP per capita, we consider two migration episodes that are examples for the largest migration waves to the US: the migration of Norwegians in the 1880s, and Mexicans in the 2000s.

Our research design combines non-parametric estimation with a calibrated simulation exercise. Based on panel data, we first estimate the degree of self-selection for both countries, confirming previous findings from the literature that Norwegians were on average mildly positively, and Mexicans negatively selected from the respective population (Fernández-Huertas Moraga, 2011; Abramitzky *et al.*, 2012). To quantify the aggregate effect of migrant self-selection, we conduct the following thought experiment: we send all migrants back to their home country, and replace them with the same number drawn at random from the full population. Put differently, we compare aggregate outcomes under the observed selective migration flows with a counterfactual world in which migrants have the same characteristics as all nationals of the sending country.

Our findings demonstrate that migrant self-selection can — but does not necessarily have to — matter at the aggregate level. Indeed, it only affects GDP per capita if two conditions are met. First, the number of migrants has to be sufficiently large relative to the full population, and second, migrants have to be sufficiently different from the average person in their home country. Among the four cases studied in this paper, only Mexico meets both conditions, and we find a large effect of migrant selection on GDP per capita. The strong negative selection of Mexican emigrants leaves the country with a more productive workforce, leading to a GDP per capita that is more than 1% higher than it would be if Mexican emigrants had the same skills as the average Mexican. This effect is as large as the aggregate impact of migration *per se* — the difference between zero migration and the current level of migration. By contrast, we find much smaller effects in Norway. While Norway had the same share of emigrants as Mexico, its emigrants had similar skills as the average Norwegian, resulting in an impact on GDP per capita of -0.26%.

The literature often motivates studying self-selection by its potentially large welfare effects in the receiving countries. Our results suggest that the impact on the receiving countries is – at best – modest. The fact that 10 million Mexican immigrants are negatively and not neutrally selected decreases GDP per capita in the US by a mere 0.28%. Two stylized facts can explain why this effect is not larger. First, the US population is larger than the population in Mexico, and thus the relative size of the migration flow is smaller in the US. Second, human capital transferability from Mexico to the US is low, such that the skill dispersion of migrants in the Mexico is wider than in the US. Mexican immigrants in the US are so heavily concentrated at the lower end of the US skill distribution that even a significant change in migrant selection only results in a modest effect. Moreover, even if we assume that all immigrants in the US are selected like Norwegians or Mexicans, the impact on per-capita income remains fairly low, at $+0.14\%$ and -0.94%, respectively.

In light of these findings, researchers need to be careful when claiming that selection has significant welfare impacts. More often than not, the conditions for observing a large welfare impact are not met. Among the four cases presented in this paper — Mexico, Norway, and the US in 1880 and the 2000s — selection has an impact close to zero in three of them. Studying self-selection of migrants might be a worthwhile demographic exercise, and could enhance our understanding of migration processes in general, although we have to be careful — not to say selective — in terms of claiming that it has broader economic consequences.

References

Abramitzky, Ran, Boustan, Leah Platt, & Eriksson, Katherine. 2012. Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration. *American Economic Review*, **102**(5), 1832–1856.

Akee, Randall K. Q. 2007. Who Leaves and Who Returns? Deciphering Immigrant Self-Selection from a Developing Country. *IZA Discussion Paper*, **3268**.

Ambrosini, J. William, & Peri, Giovanni. 2012. The Determinants and the Selection of Mexico-US Migrants. *The World Economy*, **35**(2), 111–151.

Ambrosini, J. William, Mayr, Karin, Peri, Giovanni, & Radu, Dragos. 2011. The Selection of Migrants and Returnees: Evidence from Romania and Implications. *NBER Working Paper*, **16912**.

Atack, Jeremy, Bateman, Fred, & Gregson, Mary Eschelbach. 1992. Matchmaker, Matchmaker, Make Me a Match: A General Personal Computer-Based Matching Program for Historical Research. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, **25**(2), 53–65.

Bartolucci, Cristian, Villosio, Claudia, & Wagner, Mathis. 2013. Who Migrates and Why? *Collegio Carlo Alberto, mimeo.*

Belot, Michèle V. K., & Hatton, Timothy J. 2012. Immigrant Selection in the OECD. *The Scandinavian Journal of Economics*, **114**(4), 1105–1128.

Bertoli, Simone. 2010. The Informational Structure of Migration Decision and Migrants Selection. *Economics Letters*, **108**, 89–92.

Bertoli, Simone, & Rapoport, Hillel. 2013. Heaven's Swing Door: Endogenous Skills, Migration Networks and the Effectiveness of Quality-Selective Immigration Policies. *IZA Discussion Paper*, **7749**.

Bertoli, Simone, Fernandez-Huertas Moraga, Jesus, & Ortega, Francesc. 2010. Crossing the Border: Self-Selection, Earnings and Individual Migration Decisions. *IZA Discussion Paper*, **4957**.

Biavaschi, Costanza. 2012. Recovering the Counterfactual Wage Distribution with Selective Return Migration. *IZA Discussion Paper*, **6795**.

Borjas, George J. 1987. Self-Selection and the Earnings of Immigrants. *The American Economic Review*, **77**(4), 531–553.

Borjas, George J., & Bratsberg, Bernt. 1996. Who Leaves? The Outmigration of the Foreign-Born. *The Review of Economics and Statistics*, **78**(1), 165–176.

Borjas, George J., Bronars, Stephen G., & Trejo, Stephen J. 1992. Self-Selection and Internal Migration in the United States. *Journal of Urban Economics*, **32**(2), 159–185.

Borjas, George J., Grogger, Jeffrey, & Hanson, Gordon H. 2008. Imperfect Substitution between Immigrants and Natives: a Reappraisal. *NBER Working Paper*, **13887**.

Borjas, George J., Kauppinen, Ilpo, & Poutvaara, Panu. 2013. Self-Selection of Emigrants from a Welfare State. *University of Munich, mimeo.*

Brücker, Herbert, & Trübswetter, Parvati. 2007. Do the Best Go West? An Analysis of the Self-Selection of Employed East-West Migrants in Germany. *Empirica*, **34**, 371–395.

Caponi, Vincenzo. 2010. Heterogeneous Human Capital and Migration: Who Migrates from Mexico to the US? *Annales d'Economie et de Statistique*, **97/98**, 207–234.

Carrington, William J., Detragiache, Enrica, & Vishwanath, Tara. 1996. Migration with Endogenous Moving Costs. *The American Economic Review*, **86**(4), 909–930.

Caselli, Francesco, & Coleman, Wilbur John. 2006. The World Technology Frontier. *The American Economic Review*, **96**(3), 499–522.

Chiquiar, Daniel, & Hanson, Gordon H. 2005. International Migration, Self-Selection and the Distribution of Wages: Evidence from Mexico and the United States. *Journal of Political Economy*, **113**(2).

Chiswick, Barry R. 1999. Are Immigrants Favorably Self-Selected. *The American Economic Review Papers and Proceedings*, **89**(2), 181–185.

de Coulon, Augustin, & Piracha, Matloob. 2005. Self-Selection and the Performance of Return Migrants: the Source Country Perspective. *Journal of Population Economics*, **18**, 779–807.

Di Giovanni, Julian, Levchenko, Andrei A., & Ortega, Francesc. 2012. A Global View of Cross-Border Migration. *IZA Discussion Paper*, **6584**.

DiNardo, John, Fortin, Nicole M., & Lemieux, Thomas. 1996. Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica*, **64**(5), 1001–1044.

Docquier, Frédéric, Machado, J oel, & Sekkat, Khalid. 2012. Efficiency Gains from Liberalizing Labor Mobility. *IRES Discussion Paper*, **23**.

Dustmann, Christian, Frattini, Tommaso, & Preston, Ian P. 2013. The Effect of Immigration along the Distribution of Wages. *The Review of Economic Studies*, **80**, 145–173.

Felbermayr, Gabriel J., & Kohler, Wilhelm. 2007. Immigration and Native Welfare. *International Economic Review*, **48**(3), 731–760.

Fernández-Huertas Moraga, Jesús. 2011. New Evidence on Emigrant Selection. *The Review of Economics and Statistics*, **93**(1), 72–96.

Fernández-Huertas Moraga, Jesús. 2013. Understanding Different Migrant Selection Patterns in Rural and Urban Mexico. *Journal of Development Economics*, **forthcoming**.

Ferrie, Joseph P. 1996. A new sample of males linked from the public use microdata sample of the 1850 US federal census of population to the 1860 US federal census manuscript schedules. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, **29**(4), 141–156.

Goldin, Claudia, & Katz, Lawrence F. 2007. Long-Run Changes in the Wage Structure: Narrowing, Widening, Polarizing. *Brookings Papers on Economic Activity*, **38**(2), 135–165.

Hamilton, Bob, & Whalley, John. 1984. Efficiency and Distributional Implications of Global Restrictions on Labour Mobility: Calculations and Policy Implications. *Journal of Development Economics*, **14**(1-2), 61–75.

Hatton, Timothy J, & Williamson, Jeffrey G. 1998. *The age of mass migration: Causes and economic impact.* Oxford University Press.

Ibarraran, Pablo, & Lubotsky, Darren. 2007. Mexican Immigration and Self-Selection: New Evidence from the 2000 Mexican Census. *Chap. 5, pages 159 – 192 of:* Borjas, George J. (ed), *Mexican Immigration to the United States.* University Of Chicago Press.

Iranzo, Susana, & Peri, Giovanni. 2009. Migration and Trade: Theory with an Application to the Eastern-Western European Integration. *Journal of International Economics*, **79**, 1–19.

Jensen, Adolph. 1931. Migration Statistics of Denmark, Norway and Sweden. *Pages 283–312 of: International Migrations, Volume II: Interpretations*. NBER.

Kaestner, Robert, & Malamud, Ofer. 2013. Self-Selection and International Migration: Evidence from Mexico. *Review of Economics & Statistics*, **forthcoming**.

Kanbur, Ravi, & Rapoport, Hillel. 2005. Migration Selectivity and the Evolution of Spatial Inequality. *Journal of Economic Geography*, **5**, 43–57.

Kennan, John. 2013. Open Borders. *Review of Economic Dynamics*, **16**(2), L1–L13.

Klein, Paul, & Ventura, Gustavo. 2007. TFP Differences and the Aggregate Effects of Labor Mobility in the Long Run. *B.E. Journals in Macroeconomics*, **7**(1).

Klein, Paul, & Ventura, Gustavo. 2009. Productivity Differences and the Dynamics of International Labour Movements. *Journal of Monetary Economics*, **56**, 1059–1073.

Lacuesta, Aitor. 2010. A Revision of the Self-selection of Migrants Using Returning Migrant's Earnings. *Annales d'Economie et de Statistique*, **97/98**, 235–259.

Lucas, Robert E. 1988. On the Mechanics of Economic Development. *Journal of Monetary Economics*, **22**, 3–42.

McKenzie, David, & Rapoport, Hillel. 2010. Self-Selection Patterns in Mexico-U.S. Migration: The Role of Migration Networks. *Review of Economics and Statistics*, **92**(4), 811–821.

McKenzie, David, Stillman, Steven, & Gibson, John. 2010. How Important is Selection? Experimental vs. Non-Experimental Measures of the Income Gains from Migration. *Journal of the European Economic Association*, **8**(4), 913–945.

Minnesota Population Center. 2008 (Minneapolis: Minnesota Population Center,). *North Atlantic Population Project: Complete Count Microdata*. Version 2.0 [Machine-readable database].

Orrenius, Pia M., & Zavodny, Madeline. 2005. Self-Selection among Undocumented Immigrants from Mexico. *Journal of Development Economics*, **78**, 215–240.

Ottaviano, Gianmarco, & Peri, Giovanni. 2012. Rethinking the Effects of Immigration on Wages. *Journal of the European Economic Association*, **10**(1), 152–197.

Pedersen, Peder J., Pytlikova, Mariola, & Smith, Nina. 2008. Selection or Network Effects? Migration Flows into 27 OECD Countries, 1990-2000. *European Economic Review*, **52**, 1160–1186.

Piracha, Matloob, & Vadean, Florin. 2013. Migrant educational mismatch and the labour market. *In:* Constant, Amelie F., & Zimmermann, Klaus F. (eds), *International Handbook of the Economics of Migration*. Edward Elgar.

Rooth, Dan-Olof, & Saarela, Jan. 2007. Selection in Migration and Return Migration: Evidence from Micro Data. *Economics Letters*, **94**(1), 90–95.

Rosso, Anna. 2014. Understanding Migrant Selection: the Case of Poland. *University College London, mimeo.*

van Leeuwen, Marco, Maas, Ineke, & Miles, Andrew. 2002. *HISCO - Historical International Standard Classification of Occupations*. Leuven University Press.

Williamson, Jeffrey G. 1995. The Evolution of Global Labor Markets since 1830: Background Evidence and Hypotheses. *Explorations in Economic History*, **32**, 141–196.

Xing, Chunbing. 2010. Migration, Self-Selection and Income Distributions: Evidence from Rural and Urban China. *IZA Discussion Paper*, **4979**.

Yeaple, Stephen Ross. 2005. A Simple Model of Firm Heterogeneity, International Trade, and Wages. *Journal of International Economics*, **65**, 1–20.

APPENDIX

A Data and variable construction

Section 4 briefly introduced how our selection measures were constructed. We present in this section the details of our analysis.

Matching between the 1865 Norwegian and 1880 US Censuses. We start from the 1865 Norwegian Census, which includes a 100% sampling for all Norway. IPUMS includes a total of 1,684,480 person records. We focus on individuals aged 15-40 years old in 1865 and match them forward with the Norwegian-born in the 1880 US Census.

Our matching procedure is based on a unique combination of name, surname and age. We cannot use further matching criteria such as province of birth, as this information is not available in the US Census. Since the US Census of 1880 does not report the migrants' year of immigration, we cannot restrict the analysis to migrants who arrived between 1865 and 1880. However, given that migration flows to the US prior to 1865 were rather limited (Norwegian Statistics), most Norwegian-born men in the given age range should have arrived during the period of interest and appear in the 1865 Norwegian sample.

The matching procedure follows an iterative algorithm that has become standard in the economic history literature (Ferrie, 1996; Abramitzky *et al.*, 2012).

- We identify 317,321 Norwegian men aged 15 to 40 years old in the 1865 Census, of which 263,825 are unique by name and age combinations. We keep only the unique combinations of name and age. There are 47,413 Norwegian-born men aged 30-55 in the 1880 US Census, of which 32,784 have a unique combination of name and age.

- We standardize all first and last names using the NYSIIS algorithm (Atack *et al.*, 1992), which is a phonetic algorithm for transliterating names by their sound. Hence, names with similar pronunciation will be encoded with the same string so that matching can occur despite differences in spelling. For instance, names like Jon and John will have the same NYSIIS code JAN.

- We first match by name and exact age. If a unique match is found, the observation is considered as matched. We then proceed by matching within a one-year band around the age. We delete duplicate matches, i.e. different individuals in 1865 matched to the same individual in 1880 and multiple individuals in 1880 matched to one individual in 1865. This procedure yields to a total sample of 4,108 migrants and 245,343 non-migrants. The implied emigration rate is 1.67%.[13] The implied matching rate, measured as the share of successful matches over the share of potential matches, is about 13%. This rate is similar to that in (Ferrie, 1996), while smaller than rates in (Abramitzky *et al.*, 2012). Such difference might arise from the additional challenges in matching earlier censuses, where name errors are potentially more frequent and mortality rates higher.

To assess the possible discrepancies introduced in the matching procedure, Table 4 shows the average characteristics of the matched and unmatched observations in the US, where observable traits of all individuals can be compared. It should be noted that this sample differs from the final sample, because we could not assign income levels to about 20% of matched individuals; moreover, restrict the analysis to individuals with income within two standard deviations from the mean in Norway. However, we prefer to assess the validity on the matching procedure abstracting from the additional limitations due to the lack of occupational information in Norway[14], hence comparing the full matched sample with the non-matched sample.

Table 4: Comparison of matched and non-matched sample in the US

	Matches	Non-Matches: Not-Unique	Non-Matches: Unique not found
Log-Income/Average	0.956	1.076	0.967
	(0.924)	(0.985)	(0.938)
Urban Residence	0.174	0.186	0.161
	(0.379)	(0.389)	(0.367)
Age	40.984	39.761	40.517
	(7.439)	(7.437)	(7.517)
More than 3 Children	0.373	0.333	0.356
	(0.484)	(0.471)	(0.479)
Married	0.811	0.783	0.793
	(0.392)	(0.412)	(0.405)
	4108	14629	28676

Standard deviations in parentheses.
Sample of Norwegian men, age 30-55 in the US.

Non-matches are caused by either non-uniqueness of name-age combinations (column 2) or because unique name-age combinations were not found in Norway (column 3). Discrepancies across the samples are statistically significant but seem economically very small in magnitude. Average income is lower in the matched sample and matched migrants come from larger families, although the likelihood of living in an urban area, the average

[13] This is about 60% as large as the official emigration rate from Norwegian Statistics. Such difference could be driven by name-age doubles, inability of the matching procedure to find the migrants in the US, mortality and name changes or transcription errors. Using a mortality rate of 25% (Abramitzky *et al.*, 2012) would explain half of our match failures.

[14] Average characteristics pre-migration are reported in Table 5.

age and marital status do not differ substantially across samples. Differences between the samples do not seem large enough to imply substantial changes in the conclusions of the paper.

As mentioned in the main text, we assign income levels to each occupation held by the matched sample. Income levels are available for 79.29% of the original sample. We use the log of occupation-based mean earnings in deviation from the mean of the year throughout the analysis. Furthermore, in constructing the shares of individuals in each decile of the distribution, we restrict the sample to earnings within two standard deviations from the mean, so to prevent a large mass of individuals falling in only a few deciles. Average pre-migration characteristics of the final sample are reported in Table 5.

Mexico For Mexico, we follow the sample construction and definitions reported in Fernández-Huertas Moraga (2011). We keep individuals 25-65 years old. Observations in our analysis are person-quarters. Wages are constructed by bringing wages earned in the week prior the survey to the monthly level. Individuals who were not working in the reference week were dropped. Following Fernández-Huertas Moraga (2011) and Chiquiar & Hanson (2005), we further drop individuals who worked more than 84 hours or fewer than 20 hours per week, as well as the highest and lowest 0.5% of observations, to eliminate outliers. We also drop observations for people who worked in the United States (mostly border workers as reported in Fernández-Huertas Moraga, 2011). Real wages are constructed with inflation data from the OECD. These are quarterly averages based on June 2010 and brought to January 2013 with an index of 110.831. The exchange rate of 12.748 pesos per dollar, from the International Financial Statistics of the IMF, corresponds to January 1, 2013. Hourly wages are computed by dividing the monthly wage income by 4.5 times the number of hours worked in the previous week. The quarter average is computed by pooling observations for men and women and for migrants and non-migrants. Individuals are considered to live in a rural area when their locality has fewer than 2,500 inhabitants according to the 2000 Mexican Census. As for Norway, the wage distribution is constructed by restricting the sample to wages within two standard deviations from the mean.

Summary Statistics Having constructed the two datasets as explained, Table 5 shows the summary statistics of our samples.

The first three columns of Table 5 show the characteristics of migrants and non-migrants in Norway, before migration. Migrants have higher pre-migration income and are more likely to reside in cities compared to non-migrants, but exhibit similar characteristics

Table 5: Average Characteristics by Migration Status

	Source Countries					
	Norway, 1865			Mexico, 2000-2004		
Variable	Full Population	Non-Migrants	Migrants	Full Population	Non-Migrants	Migrants
Log Income/Wage relative to average	0.957	0.956	1.013	1.077	1.079	0.880
	(0.461)	(0.461)	(0.472)	(0.857)	(0.858)	(0.714)
Urban Residence	0.151	0.150	0.249	0.807	0.808	0.643
	(0.358)	(0.357)	(0.433)	(0.395)	(0.394)	(0.479)
More than 3 Children	0.095	0.095	0.088	0.050	0.049	0.092
	(0.293)	(0.293)	(0.284)	(0.217)	(0.216)	(0.290)
More than 3 Children x Urban	0.011	0.011	0.017	0.031	0.031	0.044
	(0.103)	(0.103)	(0.131)	(0.172)	(0.172)	(0.205)
Age	27.864	27.869	27.474	40.117	40.160	36.202
	(7.184)	(7.186)	(7.000)	(10.407)	(10.412)	(9.140)
Years of Education		-	-	8.255	8.264	7.389
				(4.909)	(4.912)	(4.530)
Observations	226216	223276	2940	1203709	1192232	11477
	Destination Country					
	U.S., 1880			U.S., 2000s		
	Natives	Norwegians	Others	Natives	Mexicans	Others
Log Income/Wage relative to average	0.946	0.900	1.124	0.930	0.575	0.920
	(0.412)	(0.349)	(0.392)	(0.518)	(0.354)	(0.571)
Urban Residence	0.221	0.190	0.475			
	(0.415)	(0.393)	(0.499)			
More than 3 Children	0.098	0.131	0.147	0.028	0.119	0.040
	(0.297)	(0.338)	(0.354)	(0.165)	(0.324)	(0.195)
More than 3 Children x Urban	0.011	0.014	0.069	0.821	0.801	0.811
	(0.102)	(0.117)	(0.254)	(0.384)	(0.399)	(0.392)
Age	26.517	28.949	30.215	41.676	37.008	40.888
	(6.856)	(6.561)	(6.696)	(10.071)	(9.088)	(9.999)
Years of Education				13.559	8.920	13.552
				(2.410)	(4.232)	(3.609)
Observations	7445693	53830	1734074	415893	18823	46940

Standard deviations in parentheses.
The Norwegian sample of migrants was obtained by matching the 1865 Norwegian Census with the 1880 US Census as explained in the previous section. We further restricted the analysis to individuals with available income levels and with income within two standard deviations from the mean, so to prevent a large mass of individuals falling in only a few deciles.
In Norway, the income variable measures the occupation-based mean annual income in 2013 US dollars, relative to its average. In Mexico, the income variable represents the hourly wage relative to the quarter average. This variable has been constructed as in Fernandez Huertas-Moraga (2011). See text for explanation. In the US, the income variable is constructed as the hourly wage relative to the mean of the year, in 2013 US dollars.

with respect to their age distribution and number of children.[15]

[15] Abramitzky *et al.* (2012) find that migrants 3-15 years old in 1865 who migrate by 1900 are on average positively selected, although selection is small. Substantial skill differences arise in their results when comparing urban and rural migrants. Note that we are here discussing an earlier emigration cohort (those aged 15-30 in 1865). We applied our same matching procedure to a sample based on the 1875 and 1880 censuses, where we could find the same selection pattern of the cohort in Abramitzky et al. The results for our sample do not seem driven by the different estimation

The second part of the panel shows the same characteristics for the Mexican sample. On average, migrants are negatively selected, as indicated by the lower pre-migration wages (about 20%), younger age and lower educational attainment.

We report similar statistics for the US sample. A clear pattern emerges: while both Norwegian and Mexican migrants underperform the native population and the other migrants in their respective time periods, Norwegians are doing comparatively better than the Mexican migrants.

B Calibration

We recover the preference parameters β, θ, and σ, the returns to skill g_X, and the fixed costs F_X from a minimum distance procedure, in which we match five moments generated by the model with the corresponding moments in the data. The moments are: the share of workers in sector Y, the Gini coefficient, and the first, second, and fourth quintile of the mean-centered wage distribution. For the preference parameters we choose the starting values as in Iranzo & Peri (2009), $\beta = 0.5$, $\theta = 1.5$ and $\sigma = 4$. For g_X we use as starting value the estimates from Caselli & Coleman (2006) for Mexico and the US in 2002 ($g_{X,Mex} = 2.02$, $g_{X,US2002} = 1.68$), and for Norway and the US in 1880 we follow Goldin & Katz (2007) and take the income ratio between the 90th and the 50th percentile of the earnings distribution ($g_{X,Nor} = 1.59$, $g_{X,US1880} = 1.65$). To obtain a starting value for the fixed costs, we first search for a value that matches the employment share in sector Y given the other starting values. The starting values are $F_{X,Nor} = 5,540$; $F_{X,US1880} = 20,000$; $F_{X,Mex} = 39,350$; $F_{X,US2002} = 19$. Table 6 shows the difference between the target moments and the corresponding moments generated by the calibrated model. The matched parameters can be found in Section 5 in the main body of the paper.

Table 6: Model fit: generated vs. target moments

	Norway 1865		**USA 1880**		**Mexico 2000s**		**USA 2000s**	
	target	model	target	model	target	model	target	model
Share in sector Y	0.78	0.79	0.85	0.89	0.71	0.82	0.29	0.26
Gini	0.46	0.37	0.36	0.38	0.46	0.39	0.44	0.38
1st quintile	0.42	0.64	0.71	0.61	0.51	0.67	0.42	0.68
4th quintile	1.17	1.21	1.21	1.13	1.38	1.20	1.12	1.21
5th quintile	1.43	1.42	1.62	1.43	2.50	1.46	1.73	1.46

Note: This table shows the target moments and the corresponding moments generated by the model based on the calibrated. Wage quintiles are calculated from deciles of the wage distribution. The wage quintiles are centered to the mean.

strategy adopted in this paper and instead seem to suggest that earlier migration cohorts were more positively selected than later ones.

C Numerical procedure to find $\bar{Z}$

At the core of the general equilibrium model described in Section 3 lies the cutoff skill level $\bar{Z}$ between the traditional sector Y and the more advanced sector X. $\bar{Z}$ is determined endogenously, and depends on the structural parameters of the model and the skill distribution of the workforce. With a change in the migrant selection pattern, the skill distribution changes, which leads to a re-allocation of skill types between the two sectors. As shown by Iranzo & Peri (2009), $\bar{Z}$ is defined by the implicit function[16]

$$\begin{aligned}
\Psi(\bar{Z}, g_Y, g_X, \beta, \sigma, \theta, M, F_X, G(Z)) &= \int_0^{\bar{Z}} \exp(g_Y Z) dG(Z) \\
&- \left(\frac{1-\beta}{\beta}\right)^{\theta} \left(\frac{\sigma}{\sigma-1}\right)^{\theta-1} \left(\frac{\sigma F_X}{M\Lambda}\right) \\
&\times \frac{\exp(\theta g_Y \bar{Z})}{\exp(\theta g_X \bar{Z})} \left(\int_{\bar{Z}}^1 \exp(g_X Z) dG(Z)\right)^{\frac{\sigma-\theta}{\sigma-1}} \\
&= 0. \qquad (10)
\end{aligned}$$

For every country and every migration scenario, we compute a different value for $\bar{Z}$. Once we know $\bar{Z}$, we can re-calculate the unit costs C_X, which are the intercept of the equilibrium wage schedule for sector X shown in Figure 2. C_X, in turn, allows us to compute the equilibrium wage schedule using Equation (2), the average nominal wage using Equation (3), and the price index P_X.

In theory, the skill level Z is a continuous variable, with a continuous probability density function $g(Z)$. However, to recover $g(Z)$ from wage data, we face a trade-off between the bin size of the density function and the precision of the estimates. A smaller bin size translates into a smoother distribution, but is estimated with lower precision. Making the bin size infinitely small, we would have a truly continuous function, although its estimation would be impossible. As a solution to this trade-off, we construct $g(Z)$ using 10 bins over the support of two standard deviations above and below the mean income. Restricting the distribution to these brackets around the mean cleans the distribution from the influence of outliers. Otherwise, there would be an extreme concentration of the probability mass around the mean, and very little mass in the tails.

Having a stepwise density function means that we may not find an exact solution

[16] Recall that g_Y and g_X are the returns to skill in the respective sector, β is the weight of good X in a CES utility function, σ is the elasticity of substitution between X and Y, θ is the elasticity of substitution between varieties of good X, M is the size of the native population in the receiving countries and the full population in the sending countries, F_X is the fixed cost of production in the advanced sector.

to Equation (10), but rather an optimal $\bar{Z}$ that lies within one of the bins of the skill distribution. To find the exact value for $\bar{Z}$, we proceed as follows.

1. For every decile of the stepwise skill distribution, we compute the value of the implicit function $\Psi(\cdot)$, using the upper bound of every bin as $\bar{Z}$.
2. Using the function values for step 1, we approximate Ψ by a fourth-order polynomial, and determine the exact cutoff $\bar{Z}$ numerically, using $Z = 0.5$ as initial guess.
3. Let n be the bin that contains $\bar{Z}$. Using Equation (2), bins $1, ..., n-1$ are assigned the wage for sector Y, and $n+1, ..., 10$ the wage for sector X. The wage in bin n is a weighted average between both wages. As an example, let $\bar{Z} = 0.37$, which means that it lies in the 4th decile of the skill distribution. The weight of wage Y in bin n would then be 0.7.

D Gini-coefficient

We calculate the Gini index based on real wages according to the formula:

$$gini = 1 - \frac{\sum_{i=1}^{10} g(Z)(S_{i-1} + S_i)}{S_{10}}, \tag{11}$$

with $S_n = \sum_{i=1}^{n} g_i(Z)W_i(Z)$. $W_i(Z)$ is the wage of the i-th decile of the skill distribution. $g_i(Z)$ is the i-th decile of the skill distribution.

E The model with trade

In Section 5.3.2 we extend the basic model and allow for costless trade between both countries. When goods are traded costlessly, people in both countries have access to all varieties of good X. In equilibrium, the price for good Y, the composite price for good X and the aggregate price index are the same in both countries. The cutoff skill-levels between sectors Y and X in countries 1 and 2, $\bar{Z}_1^T$ and $\bar{Z}_2^T$ are now jointly determined in equilibrium. As before, wages between sectors Y and X have to be equal at the cutoff, such that the marginal worker is indifferent in equilibrium, $W_{Yj}(\bar{Z}_j^T) = W_{Xj}(\bar{Z}_j^T)$. The equilibrium cutoffs are pinned down by the market clearing conditions for goods Y and X,

$$
\begin{aligned}
(1 \quad - \quad & s(P_X))(M_1\bar{W}_1 + M_2\bar{W}_1) \qquad (12) \\
\quad = \quad & M_1\Lambda_1 \int_0^{\bar{Z}_1^T} \exp(g_{Y1}Z)dG_1(Z) + M_2\Lambda_2 \int_0^{\bar{Z}_2^T} \exp(g_{Y2}Z)dG_2(Z),
\end{aligned}
$$

$$
\frac{\exp\left[\left(g_{X1} - g_{Y1}\right)\bar{Z}_1^T\right]}{\exp\left[\left(g_{X2} - g_{Y2}\right)\bar{Z}_2^T\right]} = \left(\frac{F_{X1}}{F_{X2}}\right)^{\frac{1}{\sigma}}. \qquad (13)
$$

In the following, we derive these two conditions. Costless trade leads to a convergence in prices, so that $P_{Y1} = P_{Y2} = 1$, $P_{X1} = P_{X2} = P_X$, and $P_1 = P_2 = P$. Given that consumers have identical preferences and prices are equal across countries, the expenditure share of good X is also the same, $s(P_{X1}) = s(P_{X2}) = s(P_X)$.

World demand for good Y is then

$$
Y^d_{world} = (1 - s(P_X))(M_1\bar{W}_1 + M_2\bar{W}_2), \qquad (14)
$$

while world supply of good Y is the sum of the production in both countries

$$
Y^s_{world} = M_1\Lambda_1 \int_0^{\bar{Z}_1^T} \exp(g_{Y1}Z)dG_1(Z) + M_2\Lambda_2 \int_0^{\bar{Z}_2^T} \exp(g_{Y2}Z)dG_2(Z). \qquad (15)
$$

In equilibrium, equations (14) and (15) are equal, which yields Equation (12).
For the differentiated good X the value of demand for varieties produced in a given country has to equal the value of production in this country. The demand consists of demand from consumers at home, x_{11}, and abroad, x_{12}. Let the price for a given variety produced in country 1 be p_1; the total production of the variety in country 1 is x_1. The basic market clearing condition is then $p_1x_{11} + p_1x_{12} = p_1x_1$. Noting that the demands x_{11} and x_{12} are

$$
\begin{aligned}
x_{11} &= \frac{s(P_X)}{P_X}\left(\frac{p_1}{P_X}\right)^{-\sigma} M_1\bar{W}_1 \\
x_{12} &= \frac{s(P_X)}{P_X}\left(\frac{p_1}{P_X}\right)^{-\sigma} M_2\bar{W}_2,
\end{aligned}
$$

we obtain the demand for a variety of good X in country 1,

$$
x_1^d = p_1^{-\sigma}\frac{s(P_X)}{P_X^{1-\sigma}}(M_1\bar{W}_1 + M_2\bar{W}_2). \qquad (16)
$$

The supply of x_1 is determined by the zero profit condition, and is $x_1^s = (\sigma - 1)F_{X1}$. The

condition for country 2 can be derived analogously. Dividing the market clearing conditions for both countries by each other, and noting that $p_j = \exp\left[\left(g_{Yj} - g_{Xj}\right)\bar{Z}_1^T\right]\left(\frac{\sigma}{\sigma-1}\right)$ we obtain Equation (13). Based on $\bar{Z}_1^T$ and $\bar{Z}_2^T$, we can compute equilibrium prices and wages. As in the autarky case, nominal wages are calculated using Equation (3).
The aggregate price for good X is now

$$P_X = \left(N_1 p_1^{1-\sigma} + N_2 p_2^{1-\sigma}\right)^{\frac{1}{1-\sigma}}, \tag{17}$$

with N_j being the number of varieties of X produced in country j. It can be shown that

$$N_j = \frac{M_j \Lambda_j}{\sigma F_{Xj}} \int_{\bar{Z}_j^T}^{1} \exp(g_{Xj} Z) dG_j(Z), \tag{18}$$

so that

$$P_X = \frac{\sigma}{\sigma-1}\left(\frac{M_1\Lambda_1}{\sigma F_{X1}} C_{X1}^{1-\sigma} \int_{\bar{Z}_j^T}^{1} \exp(g_{X1}Z)dG_1(Z) + \frac{M_2\Lambda_2}{\sigma F_{X2}} C_{X2}^{1-\sigma} \int_{\bar{Z}_j^T}^{1} \exp(g_{X2}Z)dG_2(Z)\right)^{\frac{1}{1-\sigma}}. \tag{19}$$