CESifo WORKING PAPERS

10588
2023

July 2023

Why Is the Roy-Borjas Model Unable to Predict International Migrant Selection on Education? Evidence from Urban and Rural Mexico

Stefan Leopold, Jens Ruhose, Simon Wiederhold

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 10588*

Why Is the Roy-Borjas Model Unable to Predict International Migrant Selection on Education? Evidence from Urban and Rural Mexico

Abstract

The Roy-Borjas model predicts that international migrants are less educated than nonmigrants because the returns to education are generally higher in developing (migrant-sending) than in developed (migrant-receiving) countries. However, empirical evidence often shows the opposite. Using the case of Mexico-U.S. migration, we show that this inconsistency between predictions and empirical evidence can be resolved when the human capital of migrants is assessed using a two-dimensional measure of occupational skills rather than by educational attainment. Thus, focusing on a single skill dimension when investigating migrant selection can lead to misleading conclusions about the underlying economic incentives and behavioural models of migration.

JEL-Codes: F220, O150, J610, J240.

Keywords: international migration, selection, occupational skills, education.

Stefan Leopold
Kiel University / Germany
leopold@economics.uni-kiel.de

Jens Ruhose
Kiel University / Germany
ruhose@economics.uni-kiel.de

Simon Wiederhold
Halle Institute for Economic Research (IWH)
Halle University / Germany
simon.wiederhold@iwh-halle.de

July 24, 2023
We thank the seminar participants in the Applied Microeconomics seminar at Kiel University and the ESPE conference in Belgrade for the helpful comments and suggestions. We are also grateful to Jesús Fernández-Huertas Moraga for sharing his code for cleaning the ENET data.

1 Introduction

There is compelling empirical evidence that the international migrants of almost all countries have higher education levels than nonmigrants (Docquier & Marfouk, 2006; Grogger & Hanson, 2011). However, there is also an ongoing discussion regarding the interpretation of this pattern. The workhorse model used to explain international migration and migrant selection is the Roy-Borjas model (Roy, 1951; Borjas, 1987), which predicts that cross-country differences in returns to skills determine self-selection. Assuming that educational attainment is a sufficient measure of individual skills, as most of the literature assumes, differential returns to schooling (i.e., education-specific differences in *log* wages) between the destination and the origin country should be sufficient to explain migrant selection. However, the returns to one additional year of schooling are estimated to be higher in most migrant-sending (developing) countries than in most migrant-receiving (developed) countries (Psacharopoulos & Patrinos, 2018). Thus, the Roy-Borjas model would predict a negative selection of international migrants, while in reality, that selection is positive. In contrast, the Roy-Borjas model correctly predicts the negative educational selection of Mexican migrants to the United States, which is the largest and most-studied international migration flow between any two countries worldwide.[1] However, this average selection pattern conceals the fact that Mexican migrants from rural regions (i.e., localities with 2,500 inhabitants or less) are in fact *positively* selected (Fernández-Huertas Moraga, 2013). Thus, the case of Mexico-U.S. migration adds to the inconsistency between theory and the empirical evidence derived from international studies, questioning the applicability of the Roy-Borjas model in explaining international migrant selection.

In this paper, we provide evidence that previous inconsistencies regarding the role of the Roy-Borjas model in explaining migrant selection can be due to the measurement of migrant skills.[2] Differences in the returns to skills consistently predict migration between rural Mexico and the United States when skills are measured by *occupational skills*, i.e., the human capital acquired through performing job tasks, rather than through educational attainment. A key to our analysis is the use of occupational skill measures that are directly comparable between countries, as developed in Patt et al. (2021). We thus focus on Mexico-U.S. migration, as both countries run directly comparable large-scale surveys of worker skills. In fact, our analysis would not be feasible in a larger international setting because comprehensive information regarding worker skills beyond that attested to by educational attainment is not typically available. The Mexico-U.S. setting has

[1] See, e.g., Ambrosini & Peri (2012); Kaestner & Malamud (2014); Fernández-Huertas Moraga (2011, 2013); Patt et al. (2021).

[2] Previous literature has argued that the observed pattern of positive educational selection is driven either by differential wage levels (i.e., education-specific differences in *absolute* wages) (Grogger & Hanson, 2011) or education-specific migration costs (Belot & Hatton, 2012; Fernández-Huertas Moraga, 2013). We show that migrants' skills remain relevant when accounting for these potential explanations.

three additional advantages for our analysis. First, the United States is virtually the only destination country for Mexican migrants, so we can be confident that only the returns to skills in the United States (relative to those in Mexico) are relevant for Mexican migration decisions. Second, the comparison between migration from rural and urban Mexican regions is not confounded by cross-country differences in migration policies and other country-level migration barriers, which are likely important in other international settings. Finally, rural Mexico is comparable to many developing countries in terms of their low average educational level (Psacharopoulos & Patrinos, 2018), which extends the generalizability of our results to other migration flows.

We use measures of occupational skills that we have developed in earlier work (Patt et al., 2021). These skill measures are based on a representative worker survey (survey of the Consejo Nacional de Normalización y Certificación de Competencias Laborales; CONOCER survey), fielded by the Mexican government. The survey collected comprehensive information about the skills required across the many occupations of Mexico. The structure of the survey is comparable to that of the US O*NET, which allows us to construct occupational skill scores that are directly comparable across borders. We characterize the occupational skill spaces in Mexico and the United States in two primary dimensions: cognitive and manual. Cognitive occupational skills include problem-solving abilities, proactivity, and creativity, among other factors, while manual occupational skills include physical strength, proficiency in machine and tool operation, etc.

We identify migrants thorugh data from the Encuesta Nacional de Empleo Trimestral (ENET), Mexico's quarterly labor force survey. ENET surveys Mexican households over five consecutive quarters, eliciting sociodemographic information, such as age, gender, educational attainment, occupation, and earnings. Importantly, the panel structure of the survey allows the identification of migrant characteristics quartering the quarter preceding the move. We link our occupational skill measures to the worker data through detailed information on respondent occupations in ENET. To assess the potential relevance of migration costs in explaining migrant selection, our analysis also incorporates migration networks (McKenzie & Rapoport, 2010; Fernández-Huertas Moraga, 2013), wealth constraints (Fernández-Huertas Moraga, 2013), and the travel distance to the U.S. border (Patt et al., 2021). Our analysis focuses on Mexican males, as females often do not participate in the labor market in Mexico (Kaestner & Malamud, 2014).

We find positive migrant selection by education in rural Mexico and negative selection in urban Mexico, replicating the results in Fernández-Huertas Moraga (2013) based on the same data. Intriguingly, however, the selection based on occupational skills is the *same* in both regions: Mexican migrants are negatively selected on cognitive skills and positively selected on manual skills. Consistently, we find that differential returns to occupational skills between Mexico and the United States are positive predictors of migration propensity in both regions. Moreover, it is well established that Mexican migrants are negatively selected on earnings, which serve as a catch-all measure of individual productive characteristics. However, while Mexican migrants are on average negatively selected on earnings, we find the same heterogeneity between rural and urban Mexico as for the selection on education. Importantly, we can show that 72 percent of the negative earnings selection in urban regions can be explained by differential occupational and sociodemographic returns, while migration costs are not related to the earnings selection pattern. In contrast, migration costs can fully account for the positive earnings selection in rural regions, which is consistent with the idea that migration costs should hold more importance for migration from poorer, less developed regions (Belot & Hatton, 2012; McKenzie & Rapoport, 2010). Moreover, this finding implies that accounting for migration costs causes migrants and nonmigrants to be indistinguishable in terms of their premigration earnings and, thus, in terms of all relevant labor-market factors that can drive productivity differences in Mexico (including returns to occupational skills). However, since differential returns to occupational skills between rural Mexico and the United States are still a positive predictor of migration, we conclude that (perceived) returns to occupational skills in the U.S. labor market are an important factor in these migration decisions. The consistency of our findings between urban and rural Mexico supports the validity of the Roy-Borjas model in analyzing migrant behavior when migrant human capital is characterized by occupational skills rather than by educational attainment.

Overall, our results suggest caution against relying solely on educational attainment to explain migrant behavior.[3] When migrant human capital is measured by the skills that are actually required at the workplace, which are thus more likely to be the basis for migrants' returns-to-skills calculations, the Roy-Borjas model delivers consistent predictions regarding migrant selection. Consequently, studying migrant selection by level of education in the context of the Roy-Borjas model (or any other model that predicts migrant selection based on returns to skills) can result in misleading conclusions about the underlying economic incentives and behavioral models of migration.

2 Related Literature

The Roy-Borjas model (Roy, 1951; Borjas, 1987), which is the most commonly used migration model, works on the assumption that migrants should be positively selected (i.e., migrants have higher skills than nonmigrants) when the return to their skills is higher abroad than at home. Conversely, migrants should be negatively selected when the return to their skills is lower abroad than at home. A large body of literature has studied this hypothesis across different countries, time periods, and skill measures (e.g., education,

[3] For a similar argument, see Borjas (1991); Dustmann & Glitz (2011); Parey et al. (2017); Patt et al. (2021).

wages, occupations, cognitive skills, health, and age heaping) with mixed results.[4] What is arguably the most important inconsistency arises from the observation that international migrants from almost all countries have higher education levels than nonmigrants (Docquier & Marfouk, 2006; Grogger & Hanson, 2011), while the labor-market returns to education are estimated to be higher in most migrant-sending (developing) countries than they are in most migrant-receiving (developed) countries.[5]

The predictions of the Roy-Borjas model have been tested in the context of migration from Mexico to the United States, for the following three principal reasons: (i) the size and economic relevance of the migrant flow; (ii) the fact that Mexicans almost exclusively migrate to the United States; and (iii) the availability of high-quality worker data in Mexico. Evidence shows that, in contrast to the cross-country setting, Mexican migrants are negatively selected based on education (Fernández-Huertas Moraga, 2011; Kaestner & Malamud, 2014; Patt et al., 2021). This negative selection is consistent with the Roy-Borjas model because the returns to education are higher in Mexico than they are for Mexicans in the United States. However, an important study by Fernández-Huertas Moraga (2013) shows that, while Mexican migrants to the United States are negatively selected both on average and in urban regions, they are positively selected in rural regions. Such positive selection in rural regions is inconsistent with the predictions of the Roy-Borjas model.

When distinguishing between Mexican migrants who originate in different regions of Mexico, one can replicate the inconsistency seen in the Roy-Borjas model from international migration flows. In terms of educational attainment, there is a clear rural-urban divide in Mexico (see Table 1). The residents of rural Mexican areas have an average of 5.4 years of education, which is similar to the average years of education observed in developing South Asian countries (4.9 years) and Sub-Saharan African countries (5.2 years) (Psacharopoulos & Patrinos, 2018). In contrast, the average educational level of residents of urban regions in Mexico (9.3 years) is comparable to that in advanced economies (9.5 years on average).[6] Due to these striking differences between urban and rural Mexico, the selection patterns found in the Mexico-U.S. setting are potentially applicable to other international migration flows.

[4] See, among others, Abramitzky et al. (2012), Belot & Hatton (2012), Borjas et al. (2019), Chiquiar & Hanson (2005), Feliciano (2005), Gould & Moav (2016), Grogger & Hanson (2011), Ibarraran & Lubotsky (2007), Kaestner & Malamud (2014), Krieger et al. (2018), Fernández-Huertas Moraga (2011), Fernández-Huertas Moraga (2013), Orrenius & Zavodny (2005), Patt et al. (2021), Parey et al. (2017), Stolz & Baten (2012).

[5] On average, the wage return to one additional year of schooling amounts to up to 11 percent in developing countries, compared to 8 percent in developed countries (Psacharopoulos & Patrinos, 2018).

[6] If it were the case that international migration required a minimum level of educational attainment (e.g., as individuals must be capable of completing the paperwork), then the positive migrant selection in rural Mexico could simply be due to the generally low education level of the nonmigrant population. Such an explanation, however, would not hold in urban Mexico with its better educated population.

One potential explanation for the different selection patterns between urban and rural Mexico is omitted variables. Fernández-Huertas Moraga (2013) contributes the positive selection in rural Mexico to skill-specific migration costs in the form of international migrant networks and household wealth. This is in line with the research of Belot & Hatton (2012), who show that the result from Grogger & Hanson (2011), which posits that differential wage levels (i.e., education-specific differences in *absolute* wages) can explain migrant selection based on level of education, is not robust to accounting for poverty constraints.

In addition to migration costs, the measurement of migrant skills is another potential explanation for the conflicting empirical evidence on migrant selection. In fact, Patt et al. (2021) have demonstrated that the characterization of skills is important for determining the earnings potential of migrants abroad. They show that Mexican migrants to the United States are positively selected on manual occupational skills and negatively selected on cognitive occupational skills. This selection pattern is consistent with the observation that Mexican workers in the United States have higher returns to manual skills and lower returns to cognitive skills than Mexican workers in Mexico. The work of Patt et al. (2021) further suggests that migrants do *not* use their educational attainment to evaluate their earnings potential abroad, which they infer from their finding that selection on the basis of level of education no longer occurs when occupational skills are used as an alternative skill proxy in the estimation.

To summarize, the Roy-Borjas model has difficulties in explaining positive migrant selection based on level of education in both rural Mexico and across various international contexts. Although education-specific migration costs could play a role in interpreting this phenomenon, recent studies raise doubts about whether migrants assess their earning potential abroad based on their educational attainment.

3 Data and Empirical Approach

3.1 Identifying Migrants and Measuring Occupational Skills

To identify Mexican migrants, we use Mexico's quarterly labor force survey, ENET. The survey has frequently been used in previous work to study the selection of Mexican migrants (Fernández-Huertas Moraga, 2011, 2013; Patt et al., 2021). ENET is a rotating quarterly household panel that is carried out by the Mexican Instituto Nacional de Geografía y Estadística (INEGI) covering the period from Q2-2000 to Q4-2004. The survey samples a roster of nationally representative households, which are interviewed over five consecutive quarters. Migrants are identified if one or more individuals have left the household to migrate to the United States while at least one household member remained in Mexico.[7] Thus, an international migrant is defined as a person who is present in ENET in quarter t but has migrated abroad in the next quarter $t + 1$. We use information on the premigration work history (occupational information at the 4-digit level), educational attainment, labor earnings, age, gender, and household assets to characterize migrants and nonmigrants. Our analysis focuses on male workers aged 16—65.

We use detailed occupational skill measures from Patt et al. (2021) to assess the level of human capital that is acquired through the performance of tasks associated with an occupation. Patt et al. (2021) distinguish between two types of occupational skills: cognitive and manual. Cognitive skills are related to problem solving, proactivity, and creativity, whereas manual skills are related to physical skills and the use of tools. The skill measures are based on a principal component analysis (PCA) of the Mexican occupational task content as obtained from the CONOCER survey. This survey was fielded by the Mexican government in 2012 and includes over 17,250 observations; representative of the Mexican working population. CONOCER provides information on approximately 100 variables used to measure the content of jobs in 443 occupations (4-digit level). Because the U.S. worker survey O*NET contains very similar items, skill scores that are comparable across borders can be constructed. Reassuringly, Patt et al. (2021) show that separate PCAs on the Mexican CONOCER data and the U.S. O*NET data yield similar rotations on each item, which supports the idea that the structure of these surveys is indeed comparable. They use the loadings obtained from the PCA on the U.S. O*NET data to express Mexican skills in the U.S. skill metric.[8] To facilitate interpretation, Patt et al. (2021) convert the raw scores into a percentile scale based on the distribution of the scores in the 2010 U.S. Census. We merge these occupation-specific skills scores to the ENET data based on the 4-digit occupation code. We obtain individual-specific measures of occupational skills by averaging the scores over the current and all previously held occupations available in ENET during period t.

Using data from the 2010 Mexican Census, Figure 1 separately depicts the distribution of cognitive and manual occupational skills in the Mexican population for urban and rural Mexico. Two observations stand out. First, the occupational structure is surprisingly similar across both types of regions. There is a clear negative association between cognitive and manual occupational skills. The population-weighted correlation between these skills is $\rho = -0.35$ in urban areas and $\rho = -0.54$ in rural areas. Second, while an average worker in urban Mexico has a higher level of manual skills and a lower level of cognitive skills

[7] The migration of an entire household would appear as a missing observation in the ENET data. However, using data from the Mexican Family Life Survey (MxFLS), which follows households from Mexico to the United States, Kaestner & Malamud (2014) and Patt et al. (2021) show that the omission of entire-household migration does not affect the selection pattern.

[8] Denominating Mexican skills in the U.S. skill metric results from the idea that potential migrants evaluate the value of their skills in the U.S. labor market by comparing their skills to those of workers in the United States. However, the results are similar when PCA-loadings obtained from the CONOCER data are used to construct occupational skills (Patt et al., 2021).

(indicated by the red lines) than an average worker in the United States, this pattern is much more pronounced in rural Mexico. In urban Mexico, the average Mexican worker is located at the 39th percentile of the U.S. skill distribution in terms of cognitive skills and at the 59th percentile in terms of manual skills. In rural Mexico, the average Mexican worker lies at the 17th percentile in cognitive skills and at the 70th percentile in manual skills.

3.2 Migration Benefits

Our measure of migration benefits focuses on the economic motives behind migration in terms of higher earnings opportunities. Given the lack of consensus in the literature regarding whether differential returns or differential wage levels are better suited to explaining migrant selection (Borjas, 1987; Belot & Ederveen, 2011; Grogger & Hanson, 2011), we examine both differential labor-market returns and differential wage levels across skill cells. Following Ambrosini & Peri (2012) and Kaestner & Malamud (2014), we define one set of these cells by sociodemographic categories, that is, education, age, and marital status. A second set of cells is defined by occupational skill categories (Patt et al., 2021).

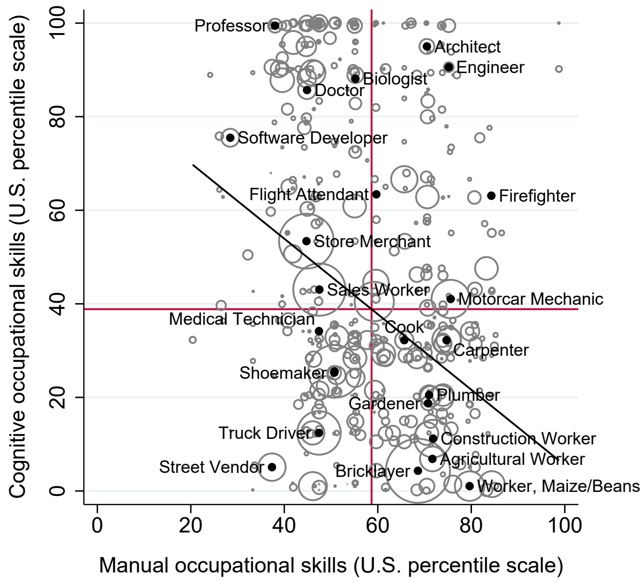
More specifically, the construction of differential labor-market returns has three steps (see Patt et al., 2021, for details). First, Mincer-type regressions are estimated separately for urban and rural regions using the 2000 Mexican Census and for Mexican migrants in the United States (who migrated to the United States between 1990 and 2000) using the 2000 U.S. Census. Thus, the monetary migration benefits are assessed before the start of the ENET survey, which avoids reverse causality concerns. The Mincer regressions include a full set of interactions between years of schooling (five categories), age (six categories), and marital status (two categories) to predict *sociodemographic returns*. For predicting *occupational returns*, the Mincer functions contain a full set of interactions between cognitive skills (four categories) and manual skills (four categories).[9] Second, based on the predicted earnings for Mexican residents and those for Mexican migrants, the labor-market returns for each skill cell are constructed by subtracting the predicted log wage of the baseline category from the predicted log wage in the respective cell. Third, the differential labor-market returns are obtained by cellwise subtraction of the labor-market return for Mexican migrants from the labor-market return for Mexican residents. Differential wage levels are calculated as the cellwise difference between the predicted wage level for Mexican migrants and the predicted wage level for Mexican residents.

The differential labor-market returns and wage levels are then merged with the ENET data using the respective categories of years of schooling, age, and marital status (sociode-

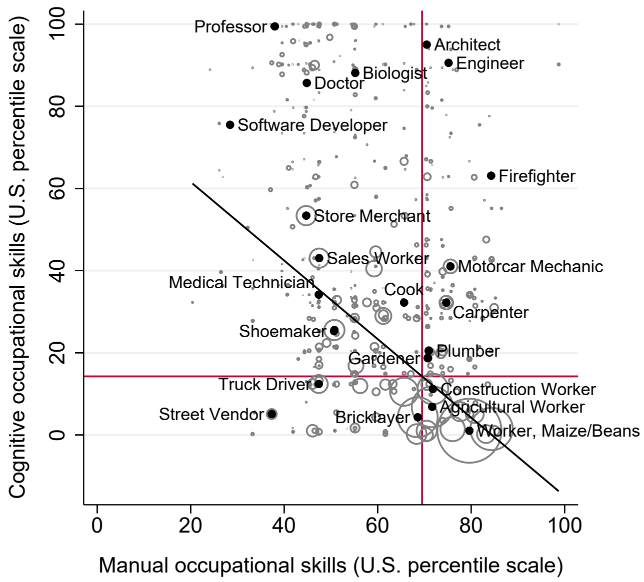
[9] The four categories for manual and cognitive skills are constructed by splitting the occupational skill distributions in the 2000 Mexican Census at their 25th, 50th, and 75th percentiles. These cutoffs are also used to construct the same skill categories from the 2000 U.S. Census.

Figure 1: Occupational Skills in the Mexican Population by Urban and Rural Status

(a) Urban areas



(b) Rural areas



Notes: The figure plots the cognitive and manual occupational skills in the Mexican population as measured in U.S. 2010 percentile ranks and weighted by the number of observations in the 2010 Mexican Census. In Figure 1a, the sample is restricted to urban areas (more than 2,500 inhabitants); in Figure 1b, the sample is restricted to rural areas (2,500 inhabitants or fewer). The sample is restricted to male Mexicans aged 16–65. Those occupations whose titles are shown are represented by solid dots; the size of the hollow circles around the solid dots is proportional to the number of Mexicans working in the occupation in the 2010 Mexican Census. The regression line (black) is weighted by the number of observations. Red lines show the weighted averages of cognitive and manual occupational skills. The population-weighted / unweighted (i.e., occupation-level) correlation between the skills is $\rho = -0.35/\rho = -0.18$ in urban areas and $\rho = -0.54/\rho = -0.19$ in rural areas. For comparison, the nationwide population-weighted/unweighted correlation between the skills is $\rho = -0.46/\rho = -0.18$. Data sources: 2010 Mexican Census, Mexican CONOCER, and U.S. O*NET.

mographic returns/wage levels) and cognitive/manual skills (occupational returns/wage levels).

3.3 Migration Costs

In terms of migration costs, we examine migration networks, household wealth, and the municipality-specific hourly travel distance to the closest U.S. border checkpoint. To construct a measure of migration networks, we follow Fernández-Huertas Moraga (2013) and McKenzie & Rapoport (2010). Both determine the migration network prevalence for a given municipality by leveraging data from the Encuesta Nacional de la Dinámica Demográfica 1997 (ENADID), which is a nationally representative demographic survey conducted by the Mexican statistical agency. The survey contains information about whether individuals from a given municipality who are older than 15 have ever been to the United States (for work or any other reasons), have moved there, or have a household member that did so. This information is collapsed at the municipality level, which provides the share of the municipality population with U.S. experience as an indicator for network prevalence. Networks are assessed before the observation period in ENET, alleviating reverse causality concerns. We merge the migration network prevalence from the ENADID with the ENET data at the municipality level.

Furthermore, some portion of workers who decide to stay in Mexico may consider it beneficial to migrate but simply cannot do so due to financial constraints. We capture financial constraints through an index of household wealth, exploiting the information on household assets that is included in ENET. Every second quarter of a given year, individuals are asked about the characteristics of their household's physical structure, such as the type of home ownership, number of rooms, and even the composition of floors and walls. Similar to Fernández-Huertas Moraga (2013), we construct the wealth index using PCA. The resulting index includes thirty-six housing characteristics and is constructed for the entire country as well as separately for urban and rural Mexico.[10]

Last, we use the travel distance in hours to the closest U.S. border checkpoint as a proxy for migration costs.

3.4 Empirical Approach

To investigate whether differential *returns* or differential wage *levels* predict migration from Mexico to the United States (Section 4.2), we estimate simple linear probability models linking differential returns or differential wage levels to migration propensity. The dependent variable, migration propensity, is a dummy that is equal to 1 if an individual migrates to the United States in the next quarter and to 0 otherwise. We scale this

[10] See Appendix Table A3 in Fernández-Huertas Moraga (2013) for all characteristics that enter the wealth index.

dummy variable by the sample-specific migrant share to make the effect sizes comparable accounting for the fact that migration probability differs substantially across quarters and regions. Thus, the coefficients are interpreted in terms of percentage changes relative to the average migration rate. We also include quarter-by-year fixed effects to control temporal migration shocks. Standard errors are clustered at the household level throughout the analysis.

Our analysis of migrant selection based on actual earnings (Section 4.3) applies linear probability models that are similar to those described above. This analysis follows the idea that actual earnings are a catch-all measure of the relevant labor-market characteristics of individuals, allowing us to assess whether individuals with different earning levels can also be associated with different types of migration benefits and migration costs.

4 Results

4.1 Descriptive Statistics

Table 1 presents the descriptive statistics of the male working population in Mexico for migrants and nonmigrants separately by type of origin region. Rural regions are defined as localities with 2,500 inhabitants or less, and urban areas are defined as having more than 2,500 inhabitants. The first row shows that migration from rural Mexico is more likely than migration from urban regions. While only 21 percent of nonmigrants live in rural regions, 43 percent of migrants to the United States originate from rural regions.

Education, wages, and other sociodemographic characteristics. Table 1 confirms the pattern of negative migrant selection with regard to education and wages that has been established in the literature (Ambrosini & Peri, 2012; Kaestner & Malamud, 2014; Fernández-Huertas Moraga, 2011, 2013). More precisely, migrants on average have 7.1 years of schooling and earn 1.6 U.S. dollars per hour (in 2006 prices, adjusted for purchasing power parity), whereas nonmigrants have 8.5 years of schooling and earn 2.2 U.S. dollars per hour.[11] However, this overall pattern of negative migrant selection conceals important differences in the selection of migrants between those from urban and those from rural regions. Migrants from urban Mexico are negatively selected, and migrants from rural Mexico are positively selected (see Fernández-Huertas Moraga, 2013, for a similar result).[12]

Occupational skills. In contrast to the inconsistent selection pattern regarding education and wages when comparing urban and rural areas, we observe the same selection pattern based on occupational skills in both types of regions. Migrants consistently have

[11] Column (4) of Appendix Table A.1 shows that negative educational selection prevails also after accounting for type of region, age, and marital status.

[12] However, positive educational selection in rural Mexico disappears when controlling for age and marital status (see Table 2, Panel B, Column (2)).

Table 1: Summary Statistics

	(1)	(2)	(3)	(4)	(5)	(6)
	Mexico		Urban Mexico		Rural Mexico	
	Non-migrants	U.S. emigrants	Non-migrants	U.S. emigrants	Non-migrants	U.S. emigrants
Living in rural area	0.207	0.439				
	(0.405)	(0.496)				
Years of schooling	8.517	7.149	9.339	8.003	5.366	6.058
	(4.764)	(3.555)	(4.640)	(3.697)	(3.826)	(3.033)
Real hourly wage	2.234	1.578	2.497	1.820	1.183	1.239
	(2.120)	(1.385)	(2.213)	(1.565)	(1.226)	(0.991)
Age	35.72	29.68	35.62	29.86	36.11	29.46
	(12.98)	(10.78)	(12.70)	(10.78)	(14.02)	(10.79)
Married	0.668	0.573	0.668	0.559	0.671	0.590
	(0.471)	(0.495)	(0.471)	(0.497)	(0.470)	(0.492)
Labor force participation	0.944	0.911	0.941	0.904	0.956	0.921
	(0.230)	(0.285)	(0.236)	(0.295)	(0.206)	(0.270)
Unemployment rate	0.0177	0.0384	0.0202	0.0518	0.00793	0.0213
	(0.132)	(0.192)	(0.141)	(0.222)	(0.0887)	(0.144)
Cognitive occupational skills	0.331	0.169	0.386	0.237	0.121	0.0814
	(0.290)	(0.199)	(0.288)	(0.220)	(0.179)	(0.121)
Manual occupational skills	0.611	0.676	0.586	0.640	0.704	0.722
	(0.132)	(0.116)	(0.126)	(0.118)	(0.112)	(0.0943)
Δ sociodemographic returns	-0.451	-0.307	-0.459	-0.325	-0.102	-0.127
	(0.360)	(0.235)	(0.330)	(0.233)	(0.215)	(0.146)
Δ occupational returns	0.152	0.316	0.0590	0.177	0.390	0.415
	(0.278)	(0.261)	(0.227)	(0.216)	(0.200)	(0.182)
Δ sociodemographic wage levels	7.858	7.745	7.714	7.596	8.434	8.164
	(0.705)	(0.722)	(0.663)	(0.667)	(0.783)	(0.835)
Δ occupational wage levels	8.385	8.090	8.353	8.014	8.397	8.295
	(1.319)	(0.904)	(1.432)	(1.062)	(0.930)	(0.729)
Network prevalence	0.0867	0.204	0.0792	0.160	0.116	0.259
	(0.100)	(0.158)	(0.0826)	(0.128)	(0.146)	(0.173)
Wealth index	-0.355	-0.890	-0.492	-0.640	-0.438	0.221
	(2.413)	(2.236)	(2.412)	(2.200)	(2.113)	(1.820)
Travel distance to border	10.24	9.836	9.823	9.429	11.83	10.35
	(5.061)	(3.999)	(5.068)	(4.338)	(4.706)	(3.448)
Observations	2,002,642	9,693	1,781,049	6,843	221,593	2,850

Notes: This table reports summary statistics for nonmigrants and U.S. emigrants by source region. Standard deviations are shown in parentheses. Urban Mexico refers to localities with more than 2,500 inhabitants. Rural Mexico refers to localities with 2,500 inhabitants or less. The sample includes Mexican males aged 16–65. See Sections 3.1 to 3.3 for further sample restrictions and variable definitions. The construction of real hourly wages follows Fernández-Huertas Moraga (2013). Wages are denoted in constant 2010 U.S. dollars, adjusted for purchasing power parity (PPP). Data sources: ENET, Mexican CONOCER, U.S. O*NET, 2000 Mexican Census (10.6% sample), and 2000 U.S. Census (5% sample).

higher average manual skills and lower average cognitive skills than Mexican residents. For instance, while nonmigrants in urban Mexico score at the 39th percentile of the U.S. cognitive skill distribution, migrants score at the 24th percentile. In terms of manual skills, nonmigrants score at the 59th percentile and migrants at the 64th percentile. The selection pattern in rural areas is similar, but the differences between migrants and nonmigrants are less pronounced. On average, nonmigrants score at the 12th percentile in cognitive skills, while migrants score at the 8th percentile. For manual skills, nonmigrants score at the 70th percentile, while migrants score at the 72nd percentile.

Table 2 provides the results from linear probability models relating the migration propensity to both types of occupational skills. For a one-decile increase in manual skills (e.g., corresponding to the manual-skill distance from a cook to a carpenter), the migration propensity increases by 20 percent in urban areas and by 6 percent in rural areas (Column (3)). In contrast, the migration propensity drops by 13.5 percent for every one-decile increase in cognitive skills (which corresponds to the cognitive-skill distance from a medical technician to a sales worker) in urban areas. The drop is somewhat less pronounced (9.7 percent) in rural areas, which is partly due to ceiling effects. In fact, as we show in Figure 1, Mexican workers in rural areas possess very high levels of manual skills and very low levels of cognitive skills, which makes it more difficult to detect substantial selection effects. The selection pattern does not change meaningfully when we control for education, age, and marital status (Columns (4) and (5) of Table 2). In fact, educational selection becomes positive in rural Mexico when controlling for occupational skills.

Migration benefits and costs. Table 1 shows that differential sociodemographic returns between the United States and Mexico are negative for both migrants and nonmigrants everywhere in Mexico. This indicates that labor-market returns to education, age, and marital status are higher in Mexico than they are for Mexicans in the United States. However, these returns are less negative for migrants than for nonmigrants, which may explain migration decisions. While the same pattern holds for migrants from urban Mexico, differential returns are more strongly negative for migrants than for nonmigrants in rural Mexico. This is inconsistent with the higher migration rates from rural areas. In contrast, differential returns to occupational skills are positive in both types of regions and are always higher for migrants than for nonmigrants. Moreover, differential sociodemographic and occupational wage levels are always positive, which is a result of the higher wage levels in the United States. However, migrants have on average lower differential wage levels than nonmigrants in both types of regions. This is again inconsistent with the economic rationale for migration.

Regarding migration costs, we observe a higher network prevalence of migrants from rural Mexico (25 percent) than those from urban Mexico (16 percent). Moreover, migrants from urban Mexico score lower on the wealth index than nonmigrants, while the opposite holds true in rural Mexico. Migrants live farther away from the U.S. border than

Table 2: Pattern of Mexican Migrant Selection: Sorted by Urban and Rural Mexico

	(1)	(2)	(3)	(4)	(5)
Dependent variable: migration propensity to the U.S.					
Panel A: Urban Mexico					
Years of schooling	-0.055***	-0.072***		0.023***	-0.008
	(0.004)	(0.005)		(0.006)	(0.006)
Age		-0.038***			-0.032***
		(0.002)			(0.002)
Married		-0.034			-0.036
		(0.056)			(0.056)
Cognitive skills			-0.135***	-0.157***	-0.119***
			(0.007)	(0.010)	(0.009)
Manual skills			0.195***	0.201***	0.192***
			(0.019)	(0.019)	(0.019)
Panel B: Rural Mexico					
Years of schooling	0.049***	-0.012		0.087***	0.024**
	(0.006)	(0.006)		(0.007)	(0.008)
Age		-0.039***			-0.037***
		(0.002)			(0.002)
Married		0.193**			0.238***
		(0.067)			(0.067)
Cognitive skills			-0.097***	-0.175***	-0.129***
			(0.014)	(0.016)	(0.016)
Manual skills			0.061*	0.058*	0.079**
			(0.026)	(0.026)	(0.026)

Notes: The sample includes Mexican males aged 16–65. The dependent variable is the migrant indicator (equal to 1 if migration to the United States occurred and 0 otherwise) scaled by the quarterly migrant share. Cognitive and manual skills incorporate fully observed premigration worker history; they are defined as the (unweighted) averages of skill content of current and all previous occupations up to four premigration quarters. Skill measures are demeaned and scaled by 10. All specifications contain quarter-by-year fixed effects. Urban Mexico ($N = 1,787,892$) refers to localities with more than 2,500 inhabitants. Rural Mexico ($N = 224,443$) refers to localities with 2,500 inhabitants or less. See Section 3.1 for sample restrictions and variable definitions. Robust standard errors, shown in parentheses, are clustered at the household level. Observations are weighted by sampling weights. Significance levels *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Data sources: ENET, CONOCER, and US O*Net.

nonmigrants, which holds across both rural and urban Mexico. However, rural migrants have a larger average travel distance than urban migrants.

4.2 Migration Benefits and Migration Propensity

In Table 3, we show the relationship between migration propensity and the different measures of the economic benefits of migration. In Column (1), we provide the results for urban Mexico when assessing migration benefits with differential returns to occupational skills and across sociodemographic characteristics.[13] Differential returns to occupational skills are significantly positively related to migration propensity. Increasing differential occupational returns by 10 percentage points increases the migration propensity by 17.4 percent. The coefficient on differential sociodemographic returns is considerably smaller, indicating an increase of only 5.5 percent in migration propensity for every 10 percentage point increase in differential returns.

Table 3: Differential Returns and Wage Levels

	Urban Mexico			Rural Mexico		
	(1)	(2)	(3)	(4)	(5)	(6)
Δ occupational returns	1.743***		1.796***	0.894***		0.928***
	(0.135)		(0.135)	(0.136)		(0.141)
Δ sociodemographic returns	0.556***		0.489***	-0.778***		-0.565***
	(0.066)		(0.072)	(0.091)		(0.104)
Δ occupational wage levels		-0.171***	-0.120***		-0.081***	-0.156***
		(0.012)	(0.012)		(0.024)	(0.026)
Δ sociodemographic wage levels		-0.296***	-0.357***		-0.442***	-0.386***
		(0.037)	(0.038)		(0.038)	(0.040)

Notes: The sample includes Mexican males aged 16–65. The dependent variable is the migrant indicator (equal to 1 if migration to the United States occurred and 0 otherwise) scaled by the quarterly migrant share. All regressions contain quarter-by-year fixed effects. Observations are weighted by sampling weights. Urban Mexico ($N = 1,787,892$) refers to localities with more than 2,500 inhabitants. Rural Mexico ($N = 224,443$) refers to localities with 2,500 inhabitants or less. See Sections 3.1 and 3.2 for sample restrictions and variable definitions. Robust standard errors, shown in parentheses, are clustered at the household level. Significance levels *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Data sources: ENET, Mexican CONOCER, U.S. O*NET, 2000 Mexican Census (10.6% sample), and 2000 U.S. Census (5% sample).

As shown in Column (4) of Table 3, the differential returns to occupational skills in rural Mexico remain positively related to migration propensity, while the coefficient on differential returns to sociodemographic returns becomes significantly *negative*. The negative coefficient implies that differential returns to sociodemographic characteristics represent migration costs in rural Mexico in the form of forgone earnings potential. That is, migrants would have been able to earn higher returns from their sociodemographic characteristics in Mexico than in the United States. This result is in line with the literature, which shows the same pattern for international migration in general (Grogger & Hanson, 2011). Thus, while migration propensity is consistently positively related to

[13] More detailed results are reported in Appendix Table A.2, where we include the variables step by step and show results for the entirety of Mexico.

differential returns to occupational skills, supporting the predictions of the Roy-Borjas model, sociodemographic returns provide an inconclusive picture of the role played by migration benefits in migration decisions.

In Columns (2) and (5) of Table 3, we analyze the differential wage levels, that is, the skill-specific differences in absolute wages. Following Grogger & Hanson (2011), we should expect a positive relationship between migration propensity and differential sociodemographic wage levels that explains the positive migrant selection based on level of education in rural Mexico because differences in wage levels between developed and developing countries are usually larger at higher education levels than they are at lower education levels. However, differential wage levels are consistently negatively related to migration propensity, both in rural and urban regions and for occupational and sociodemographic wages. The coefficients are also rather small. A 10 percentage point increase in differential occupational (sociodemographic) wage levels is associated with a reduction in migration propensity of 0.8 to 1.7 percent (3 to 4.4 percent). The results remain very similar when both types of migration benefits (returns and wage levels) are simultaneously included, as shown in Columns (3) and (6). Thus, the pattern of Mexican migrant selection cannot be explained by differential wage levels, which is consistent with the findings of Kaestner & Malamud (2014) and Patt et al. (2021). This result also raises questions about using differential wage levels to explain the selection of international migrants more generally.

4.3 Migration Benefits, Migration Costs, and Selection Based on Earnings

It is well established that Mexican migrants are negatively selected with respect to earnings (Ambrosini & Peri, 2012; Kaestner & Malamud, 2014; Fernández-Huertas Moraga, 2011; Patt et al., 2021).[14] However, Fernández-Huertas Moraga (2013) shows that the direction of earnings selection depends on region type in Mexico. Earnings selection is negative in urban areas, while it is positive in rural areas. Column (1) of Table 4 confirms this pattern. Doubling log hourly earnings is associated with a decrease in migration propensity of approximately 38 percent in urban Mexico and an increase in migration propensity of approximately 15 percent in rural Mexico.

The observed negative earnings selection in rural Mexico could be explained by a negative correlation between the benefits of migration and earnings; that is, those with the highest earnings in Mexico profit the least from migration (Kaestner & Malamud, 2014; Patt et al., 2021). However, negative earnings selection might also be explained by a positive correlation between migration costs and earnings; that is, those with the highest migration costs are those with the highest earnings. The reverse holds true for

[14] In our data, doubling the log hourly earnings is associated with a decrease in migration propensity of approximately 28 percent (Appendix Table A.4, Panel A, Column (1)).

Table 4: Selection by Earnings and Differential Returns

	(1)	(2)	(3)	(4)
Dependent variable: migration propensity to the U.S.				
Panel A: Urban Mexico				
Log hourly earnings	-0.383***	-0.107**	-0.378***	-0.182***
	(0.032)	(0.039)	(0.034)	(0.041)
Δ occupational returns		1.346***		1.178***
		(0.134)		(0.134)
Δ sociodemographic returns		0.485***		0.332***
		(0.070)		(0.072)
Network prevalence			9.652***	9.253***
			(0.638)	(0.630)
Wealth Index			0.011	0.038***
			(0.010)	(0.011)
Travel distance to US border			0.030***	0.032***
			(0.005)	(0.005)
Panel B: Rural Mexico				
Log hourly earnings	0.147***	0.167***	0.025	0.027
	(0.030)	(0.036)	(0.031)	(0.036)
Δ occupational returns		0.846***		0.724***
		(0.164)		(0.163)
Δ sociodemographic returns		-0.520***		-0.692***
		(0.112)		(0.115)
Network prevalence			5.051***	5.091***
			(0.315)	(0.316)
Wealth Index			0.049**	0.040*
			(0.015)	(0.016)
Travel distance to US border			0.005	0.004
			(0.005)	(0.005)

Notes: The sample includes Mexican males aged 16–65. The dependent variable is the migrant indicator (equal to 1 if migration to the United States occurred and 0 otherwise) scaled by the quarterly migrant share. The construction of hourly earnings follows Fernández-Huertas Moraga (2011). Hourly earnings are obtained by dividing monthly earnings by 4.5 x hours worked per week. Earnings are denoted in constant 2010 US dollars and adjusted for purchasing power parity (PPP). Earnings observations are dropped for individuals who are unemployed, not in the labor force, not working in Mexico, and who work less than 20 or more than 84 hours per week. Urban Mexico ($N = 1,360,580$) refers to localities with more than 2,500 inhabitants. Rural Mexico ($N = 162,878$) refers to localities with 2,500 inhabitants or less. See Sections 3.1 and 3.3 for sample restrictions and variable definitions. All regressions contain quarter-by-year fixed effects. Observations are weighted by sampling weights. Robust standard errors, shown in parentheses, are clustered at the household level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Data sources: ENET, Mexican CONOCER, U.S. O*NET, 2000 Mexican Census (10.6% sample), and 2000 U.S. Census (5% sample).

positive earnings selection, which might be due to those with the highest earnings in Mexico benefitting the most from migration and/or the fact that those with the highest migration costs are those with the lowest earnings.

We investigate the importance of migration benefits vis-à-vis migration costs in explaining earnings selection in Table 4.[15] In urban Mexico (Panel A), migration benefits are strongly related to earnings selection. Accounting for differential returns to both occupational skills and sociodemographic characteristics reduces the coefficient on log hourly earnings by 72 percent (Column (2)). For comparison, the conditioning on differential returns to sociodemographic returns alone accounts for approximately 50 percent of earnings selection (Appendix Table A.4, Panel B, Column (3)). This result is consistent with the finding in Fernández-Huertas Moraga (2013), who assigns approximately the same share of earnings selection to differential returns to education.[16] Thus, additionally accounting for differential returns to occupational skills substantially increases the share of earnings selection that can be explained by differential returns. Confirming the evidence in Fernández-Huertas Moraga (2013), we do not find any role played by migration costs in explaining the earnings selection in urban regions (Columns (3) and (4)).

In rural Mexico, we arrive at the opposite conclusion regarding the role of migration benefits versus migration costs in earnings selection (Panel B of Table 4). While earnings selection is unaffected by differential returns (Column (2)), it can fully be explained by migration costs (Column (3)). These findings are again in line with the evidence presented in Fernández-Huertas Moraga (2013). However, differential returns to occupational skills are still positively associated with migration propensity regardless of whether one accounts for migration costs, while returns to sociodemographic characteristics are consistently negatively related to migration propensity.

Our result showing that a worker's propensity to migrate from rural Mexico to the United States is no longer related to earnings when migration costs are accounted for indicates that migrants and nonmigrants have similar productivity levels in Mexico, have similar opportunity costs for migration (i.e., foregone earnings in Mexico) and have a similar potential to bear the direct migration costs (e.g., due to access to migrant networks and the availability of household assets). Therefore, the migration behavior of rural Mexican migrants can be explained by their perceived economic returns to occupational skills in the U.S. labor market. This conclusion supports the predictions of the Roy-Borjas model when occupational skills are used rather than sociodemographic characteristics to assess migrant skills.

15 Appendix Table A.3 shows an analogous analysis using wage levels instead of returns.

16 We find that our occupational skill measures make a very similar contribution to explaining earnings selection when including them as the only returns measure (Appendix Table A.4, Panel B, Column (2)).

5 Conclusion

We present novel evidence on the selection of international migrants from urban and rural Mexico to the United States. Our results confirm previous findings that migrants from urban Mexico are positively selected based on earnings and education, while those from rural Mexico are negatively selected based on both dimensions. Since returns to education are higher in both urban and rural Mexico than they are for Mexicans in the United States, the literature has concluded that the Roy-Borjas model is not applicable to explaining migrant selection in rural Mexico. Our study challenges this view. We show that the model is consistent with the data when we use occupational skills rather than education and other sociodemographic characteristics to assess migrants' skills.

Our paper aims to contribute to the ongoing debate concerning the applicability of the Roy-Borjas model in explaining international migration and migrant selection based on level of education. As is the case in rural Mexico, international migration is typically characterized by positive migrant selection based on level of education, which is inconsistent with the predictions of the Roy-Borjas model because returns to education are generally higher in developing (migrant-sending) countries than in developed (migrant-receiving) countries. Our findings suggest that this inconsistency could be an artifact of a more fundamental relationship between migration and occupational skills, which is only imperfectly proxied by education. However, to examine this relationship on a global scale, comparable data that combines occupational skills and migration across multiple countries would be needed. Gathering evidence on the role of occupational skills in a wider variety of migration flows would aid our understanding of the underlying economic incentives of migration and enrich the existing behavioral migration models.

References

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

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

Belot, M., & Ederveen, S. (2011, February). Cultural barriers in migration between OECD countries. *Journal of Population Economics*, *25*(3), 1077–1105.

Belot, M., & Hatton, T. J. (2012). Immigrant Selection in the OECD. *Scandinavian Journal of Economics*, *114*(4), 1105–1128.

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

Borjas, G. J. (1991). Immigration and Self-Selection. In J. M. Abowd & R. B. Freeman (Eds.), *Immigration, trade and the labor market* (p. 29-76). University of Chicago Press.

Borjas, G. J., Kauppinen, I., & Poutvaara, P. (2019). Self-selection of Emigrants: Theory and Evidence on Stochastic Dominance in Observable and Unobservable Characteristics. *Economic Journal*, *129*(617), 143-171.

Chiquiar, D., & Hanson, G. H. (2005). International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States. *Journal of Political Economy*, *113*(2), 239–281.

Docquier, F., & Marfouk, A. (2006). International migration by educational attainment (1990-2000) - release 1.1. In C. ï¿½zden & M. Schiff (Eds.), *International migration, remittances and development.* Palgrave Macmillan: New York.

Dustmann, C., & Glitz, A. (2011). Migration and Education. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (4th ed., p. 327-439). Elsevier.

Feliciano, C. (2005). Educational Selectivity in U.S. Immigration: How Do Immigrants Compare to Those Left Behind? *Demography*, *42*(1), 131–152.

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

Fernández-Huertas Moraga, J. (2013). Understanding Different Migrant Selection Patterns in Rural and Urban Mexico. *Journal of Development Economics*, *103*(1), 182–201.

Gould, E. D., & Moav, O. (2016). Does High Inequality Attract High Skilled Immigrants? *Economic Journal*, *126*(593), 1055-1091.

Grogger, J., & Hanson, G. H. (2011). Income Maximization and the Selection and Sorting of International Migrants. *Journal of Development Economics*, *95*(1), 42–57.

Ibarraran, P., & Lubotsky, D. (2007). Mexican Immigration and Self-Selection: New Evidence from the 2000 Mexican Census. In G. J. Borjas (Ed.), *Mexican immigration to the united states* (pp. 159–192). Chicago, IL: University of Chicago Press.

Kaestner, R., & Malamud, O. (2014). Self-Selection and International Migration: New Evidence from Mexico. *Review of Economics and Statistics*, *96*(1), 78–91.

Krieger, T., Renner, L., & Ruhose, J. (2018). Long-Term Relatedness between Countries and International Migrant Selection. *Journal of International Economics*, *113*, 35-54.

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

Orrenius, P. M., & Zavodny, M. (2005). Self-Selection Among Undocumented Immigrants from Mexico. *Journal of Development Economics*, *78*(1), 215–240.

Parey, M., Ruhose, J., Waldinger, F., & Netz, N. (2017). The Selection of High-Skilled Emigrants. *Review of Economics and Statistics*, *99*(5), 776-792.

Patt, A., Ruhose, J., Wiederhold, S., & Flores, M. (2021). International emigrant selection on occupational skills. *Journal of the European Economic Association*, *19*, 1249-1298. doi: 10.1093/jeea/jvaa032

Psacharopoulos, G., & Patrinos, H. A. (2018). Returns to Investment in Education: A Decennial Review of the Global Literature. *Education Economics*, *26*(5), 445-458.

Roy, A. D. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers*, *3*(2), 135-146.

Stolz, Y., & Baten, J. (2012). Brain Drain in the Age of Mass Migration: Does Relative Inequality Explain Migrant Selectivity? *Explorations in Economic History*, *49*(2), 205–220.

A Appendix

Table A.1: Pattern of Mexican Migrant Selection

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable: migration propensity to the U.S.								
Years of schooling	-0.053***		-0.018***	-0.043***			0.048***	0.011*
	(0.003)		(0.003)	(0.003)			(0.004)	(0.004)
Rural		1.292***	1.220***	1.143***		0.846***	0.900***	0.890***
		(0.060)	(0.064)	(0.063)		(0.064)	(0.065)	(0.064)
Age				-0.040***				-0.035***
				(0.001)				(0.001)
Married				0.065				0.074
				(0.042)				(0.042)
Cognitive skills					-0.134***	-0.104***	-0.150***	-0.107***
					(0.005)	(0.005)	(0.007)	(0.007)
Manual skills					0.209***	0.148***	0.158***	0.154***
					(0.013)	(0.014)	(0.014)	(0.014)

Notes: The sample includes Mexican males aged 16–65. The dependent variable is the migrant indicator (equal to 1 if migration to the United States occurred and 0 otherwise) scaled by the quarterly migrant share. Cognitive and manual skills incorporate full observed premigration worker history; they are defined as the (unweighted) averages of the skill content of all current and previous occupations for up to four premigration quarters. Skill measures are demeaned and scaled by 10. All specifications contain quarter-by-year fixed effects. See Section 3.1 for sample restrictions and variable definitions. $N = 2,012,335$. Robust standard errors, shown in parentheses, are clustered at the household level. Observations are weighted by sampling weights. Significance levels *** p < 0.01, ** p < 0.05, * p < 0.1. Data sources: ENET, CONOCER, and US O*Net.

Table A.2: Migration Propensity, Differential Returns, and Differential Wage Levels

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable: migration propensity to the U.S.							
Panel A: Mexico							
Δ occupational returns	2.004***		1.820***				1.819***
	(0.073)		(0.087)				(0.086)
Δ sociodemographic returns		1.022***	0.268***				0.355***
		(0.038)	(0.045)				(0.051)
Δ occupational wage levels				-0.158***		-0.171***	-0.107***
				(0.009)		(0.010)	(0.009)
Δ sociodemographic wage levels					-0.232***	-0.262***	-0.384***
					(0.025)	(0.026)	(0.028)
Panel B: Urban Mexico							
Δ occupational returns	2.135***		1.743***				1.796***
	(0.121)		(0.135)				(0.135)
Δ sociodemographic returns		1.140***	0.556***				0.489***
		(0.061)	(0.066)				(0.072)
Δ occupational wage levels				-0.160***		-0.171***	-0.120***
				(0.012)		(0.012)	(0.012)
Δ sociodemographic wage levels					-0.268***	-0.296***	-0.357***
					(0.036)	(0.037)	(0.038)
Panel C: Rural Mexico							
Δ occupational returns	0.669***		0.894***				0.928***
	(0.126)		(0.136)				(0.141)
Δ sociodemographic returns		-0.555***	-0.778***				-0.565***
		(0.082)	(0.091)				(0.104)
Δ occupational wage levels				-0.103***		-0.081***	-0.156***
				(0.024)		(0.024)	(0.026)
Δ sociodemographic wage levels					-0.448***	-0.442***	-0.386***
					(0.038)	(0.038)	(0.040)

Notes: The sample includes Mexican males aged 16–65. The dependent variable is the migrant indicator (equal to 1 if migration to the United States occurred and 0 otherwise) scaled by the quarterly migrant share. All regressions contain quarter-by-year fixed effects. Observations are weighted by sampling weights. Mexico N = 2,012,335, Urban Mexico $N = 1,787,892$, and rural Mexico $N = 224,443$. See Sections 3.1 and 3.2 for sample restrictions and variable definitions. Robust standard errors, shown in parentheses, are clustered at the household level. Significance levels *** p $<$ 0.01, ** p $<$ 0.05, * p $<$ 0.1. Data sources: ENET, Mexican CONOCER, U.S. O*NET, 2000 Mexican Census (10.6% sample), and 2000 U.S. Census (5% sample).

Table A.3: Selection on Earnings and Differential Wages

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: migration propensity to the U.S.						
Panel A: Mexico						
Log hourly earnings	-0.277***	-0.304***	-0.243***	-0.269***	-0.221***	-0.218***
	(0.021)	(0.022)	(0.022)	(0.022)	(0.024)	(0.025)
Δ sociodemographic wage levels		-0.273***		-0.291***		-0.354***
		(0.030)		(0.030)		(0.031)
Δ occupational wage levels			-0.105***	-0.116***		-0.082***
			(0.011)	(0.011)		(0.011)
Network prevalence					8.922***	8.986***
					(0.376)	(0.378)
Wealth Index					-0.015	-0.023*
					(0.009)	(0.009)
Travel distance to US border					0.029***	0.031***
					(0.003)	(0.003)
Panel B: Urban Mexico						
Log hourly earnings	-0.383***	-0.337***	-0.384***	-0.334***	-0.378***	-0.336***
	(0.032)	(0.033)	(0.032)	(0.033)	(0.034)	(0.035)
Δ occupational wage levels		-0.103***		-0.113***		-0.089***
		(0.014)		(0.014)		(0.014)
Δ sociodemographic wage levels			-0.270***	-0.290***		-0.309***
			(0.041)	(0.041)		(0.042)
Network prevalence					9.652***	9.644***
					(0.638)	(0.639)
Wealth Index					0.011	0.009
					(0.010)	(0.010)
Travel distance to US border					0.030***	0.031***
					(0.005)	(0.005)
Panel C: Rural Mexico						
Log hourly earnings	0.147***	0.157***	0.110***	0.119***	0.025	-0.005
	(0.030)	(0.031)	(0.030)	(0.031)	(0.031)	(0.031)
Δ occupational wage levels		-0.110***		-0.088**		-0.074**
		(0.028)		(0.028)		(0.028)
Δ sociodemographic wage levels			-0.507***	-0.499***		-0.515***
			(0.054)	(0.054)		(0.054)
Network prevalence					5.051***	5.095***
					(0.315)	(0.316)
Wealth Index					0.049**	0.052***
					(0.015)	(0.016)
Travel distance to US border					0.005	0.009
					(0.005)	(0.005)

Notes: The sample includes Mexican males aged 16–65. The dependent variable is the migrant indicator (equal to 1 if migration to the United States occurred and 0 otherwise) scaled by the quarterly migrant share. The construction of hourly earnings follows Fernández-Huertas Moraga (2011). Hourly earnings are obtained by dividing monthly earnings by 4.5 x hours worked per week. Earnings are denoted in constant 2010 U.S. dollars and adjusted for purchasing power parity (PPP). Earnings observations are dropped for individuals who are unemployed, not in the labor force, not working in Mexico, and who work less than 20 or more than 84 hours per week. See sections 3.1 and 3.3 for sample restrictions and variable definitions. All regressions contain quarter-by-year fixed effects. Observations are weighted by sampling weights. Mexico $N = 1,523,458$, Urban Mexico $N = 1,360,580$, and rural Mexico $N = 162,878$. Robust standard errors, shown in parentheses, are clustered at the household level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Data sources: ENET, Mexican CONOCER, U.S. O*NET, 2000 Mexican Census (10.6% sample), and 2000 U.S. Census (5% sample).

Table A.4: Selection on Earnings and Differential Returns

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: migration propensity to the U.S.						
Panel A: Mexico						
Log hourly earnings	-0.277***	-0.032	-0.126***	0.010	-0.221***	-0.066*
	(0.021)	(0.027)	(0.026)	(0.029)	(0.024)	(0.030)
Δ occupational returns		1.570***		1.423***		1.111***
		(0.092)		(0.096)		(0.095)
Δ sociodemographic returns			0.718***	0.308***		0.098
			(0.050)	(0.050)		(0.052)
Network prevalence					8.922***	8.537***
					(0.376)	(0.372)
Wealth Index					-0.015	0.017
					(0.009)	(0.010)
Travel distance to US border					0.029***	0.029***
					(0.003)	(0.003)
Panel B: Urban Mexico						
Log hourly earnings	-0.383***	-0.186***	-0.197***	-0.107**	-0.378***	-0.182***
	(0.032)	(0.035)	(0.038)	(0.039)	(0.034)	(0.041)
Δ occupational returns		1.581***		1.346***		1.178***
		(0.130)		(0.134)		(0.134)
Δ sociodemographic returns			0.835***	0.485***		0.332***
			(0.070)	(0.070)		(0.072)
Network prevalence					9.652***	9.253***
					(0.638)	(0.630)
Wealth Index					0.011	0.038***
					(0.010)	(0.011)
Travel distance to US border					0.030***	0.032***
					(0.005)	(0.005)
Panel C: Rural Mexico						
Log hourly earnings	0.147***	0.197***	0.120***	0.167***	0.025	0.027
	(0.030)	(0.034)	(0.034)	(0.036)	(0.031)	(0.036)
Δ occupational returns		0.731***		0.846***		0.724***
		(0.158)		(0.164)		(0.163)
Δ sociodemographic returns			-0.359***	-0.520***		-0.692***
			(0.107)	(0.112)		(0.115)
Network prevalence					5.051***	5.091***
					(0.315)	(0.316)
Wealth Index					0.049**	0.040*
					(0.015)	(0.016)
Travel distance to US border					0.005	0.004
					(0.005)	(0.005)

Notes: The sample includes Mexican males aged 16–65. The dependent variable is the migrant indicator (equal to 1 if migration to the United States occurred and 0 otherwise) scaled by the quarterly migrant share. The construction of hourly earnings follows Fernández-Huertas Moraga (2011). Hourly earnings are obtained by dividing monthly earnings by 4.5 x hours worked per week. Earnings are denoted in constant 2010 U.S. dollars and adjusted for purchasing power parity (PPP). Earnings observations are dropped for individuals who are unemployed, not in the labor force, not working in Mexico, and who work less than 20 or more than 84 hours per week. See sections 3.1 and 3.3 for sample restrictions and variable definitions. All regressions contain quarter-by-year fixed effects. Observations are weighted by sampling weights. Mexico $N = 1,523,458$, Urban Mexico $N = 1,360,580$, and rural Mexico $N = 162,878$. Robust standard errors, shown in parentheses, are clustered at the household level. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Data sources: ENET, Mexican CONOCER, U.S. O*NET, 2000 Mexican Census (10.6% sample), and 2000 U.S. Census (5% sample).