The Selection of High-Skilled Emigrants*

Matthias Parey, Jens Ruhose, Fabian Waldinger, and Nicolai Netz

Abstract

We measure selection among high-skilled emigrants from Germany using predicted earnings. Migrants to less equal countries are positively selected relative to non-migrants, while migrants to more equal countries are negatively selected, consistent with the prediction in Borjas (1987). Positive selection to less equal countries reflects university quality and grades, and negative selection to more equal countries reflects university subject and gender. Migrants to the United States are highly positively selected and concentrated in STEM fields. Our results highlight the relevance of the Borjas model for high-skilled individuals when credit constraints and

*Matthias Parey: University of Essex and Institute for Fiscal Studies, m.parey@essex.ac.uk, Jens Ruhose: Leibniz Universität Hannover, ruhose@wipol.uni-hannover.de, Fabian Waldinger: London School of Economics, f.waldinger@lse.ac.uk, Nicolai Netz: DZHW, netz@dzhw.eu. We thank the editor Amit Khandelwal and three anonymous referees for very helpful comments. We also thank Clement de Chaisemartin, Thomas Crossley, Christian Dustmann, Tim Hatton, Hans Hvide, Julian Johnsen, Gordon Kemp and participants in various seminars. We thank Kolja Briedis, Christian Kerst, and Gregor Fabian for providing access to the DZHW data. This study uses data from the Swiss Labour Force Survey (*Schweizerische Arbeitskräfteerhebung, BFS*). Parey gratefully acknowledges the support of the ESRC Research Centre on Micro-Social Change (MiSoC) at the University of Essex.

other migration barriers are unlikely to be binding. (JEL: F22, J61, O15, I23)

Introduction

International migration of high-skilled individuals has risen dramatically in recent decades (Docquier and Rapoport, 2012). Between 2000 and 2006, the United States attracted 1.9 million and European OECD countries attracted 2.2 million tertiary-educated migrants (Widmaier and Dumont, 2011). In the year 2000, high-skilled migrants represented about 11% of the tertiary-educated population in OECD countries (Brücker et al., 2012). In the United States, as of 2013, about 19% of the working-age population with a bachelor's degree or higher were foreign-born. In certain fields such as science, technology, engineering, and mathematics (STEM), more than 30% were foreign-born.[1]

Access to high-skilled individuals is central to firms' success, and has become even more important in economies where ideas drive technological progress (Chambers et al., 1998). When the homegrown pool of high-skilled individuals is insufficient, the ability to attract high-skilled migrants is crucial for improving the quality of a country's workforce and its innovative capacity. A deeper understanding of the selection of high-skilled migrants is therefore important – for sending and receiving countries alike.

While migrant selection has been studied extensively since Borjas (1987) outlined theoretical predictions for selection, few papers have studied the selection of *high-skilled* migrants. Focusing our analysis on high-skilled migrants who mostly migrate between developed countries enables us to investigate a setting where individuals face low legal barriers to migration, and relatively small migration costs. The economic forces described by the Borjas model should be particularly relevant in our setup.[2]

[1] Authors' calculations based on the 2013 ACS (Ruggles et al., 2010).

[2] See section 1.2 for a review of empirical papers investigating migrant selection across the entire skill distribution. The existing papers on migrant selection mostly focus on low-skilled migration between Mexico and the United States, where migrants face higher

A basic version of the Borjas (1987, 1991) model, building on Roy (1951), predicts that migrants to less equal countries, such as the United States, should be *positively* selected, while migrants to more equal countries, such as Denmark, should be *negatively* selected. Analyzing migration to both less *and* more equal countries is therefore particularly valuable to test the predictions of the model.

We study the selection of high-skilled emigrants by investigating migration decisions of graduates from German universities. Germany exhibits an intermediate level of inequality for high-skilled individuals (Figure 1). By studying selection to less *and* more equal countries in the same context, we can test both predictions of the Roy/Borjas model. Furthermore, we are able to test whether the predictions of the Roy/Borjas model hold within the population of university graduates.[3]

We use rich survey data on German university graduates collected by the German Centre for Higher Education Research and Science Studies (DZHW). German university-bound students represent a more selective group than their counterparts in most economically developed countries; this allows us to study migration patterns of the top 11% of the educational distribution.[4]

migration costs and legal barriers to entry. Other papers on high-skilled migrants study other outcomes, such as effects on the receiving economy (Hunt and Gauthier-Loiselle, 2010, Kerr and Lincoln, 2010, Borjas and Doran, 2012, Moser et al., 2014, Kerr et al., 2015, Doran et al., 2014) and on source countries (Docquier and Rapoport, 2012).

[3]Since many papers investigate migrant selection between two countries only (see Online Appendix Table A.1), they are limited to testing one of the two predictions of the Roy/Borjas model. While Borjas et al. (2015) study migration from Denmark to multiple destinations, they focus on positive selection because Denmark has very low levels of inequality.

[4]Administrative data show that about 11% of the cohorts that we study in our paper graduated from university. In 2012, the stock of university graduates among 35 to 44

To measure selection we compare predicted earnings of migrants and non-migrants. We first estimate an augmented Mincer regression for graduates who work in Germany. We then use the estimated returns to construct predicted earnings independently of whether the graduate stays in Germany or migrates abroad. Our data contain a rich set of personal characteristics including family background, high school grades, university education (including the specific university, subject, and grades), and information on mobility before enrolling at university. These detailed characteristics allow us to obtain predicted earnings as a precise measure of individual earnings potential, so that we can differentiate between high- and low-productivity graduates. We then compare cumulative distribution functions of predicted earnings for three groups of graduates: graduates who stay in Germany, graduates who migrate to less equal countries, and graduates who migrate to more equal countries. This allows us to investigate whether the most or least skilled university graduates stay in Germany or select into more or into less equal destinations. To classify destinations into either more or less equal countries, we construct new inequality measures for university graduates, based on individual-level income surveys from 20 countries.

The selection of university graduates is consistent with the predictions of the basic Roy/Borjas model. Migrants to *less equal* countries have significantly higher predicted earnings than non-migrants. Migrants to *more equal* countries, in contrast, have significantly lower predicted earnings than non-migrants. These findings hold along the whole distribution of predicted earnings. In fact, the selection patterns predicted by the model hold even within subgroups of either more equal or less equal countries.

The coefficients of the Mincer regression, which form the basis of our earnings prediction, might be biased if migrants were non-randomly selected from the population of graduates in a way not captured by our observed covariates. We address potential selec-

year-olds was 1,208,000 out of a population of 11,004,000 (DESTATIS, 2013, p. 27).

tion in the augmented Mincer regression using a sample selection correction (Heckman, 1979). In the selection equation we use the roll-out of ERASMUS, the largest study abroad program, as an instrumental variable to predict whether individuals move abroad or work in Germany. Changes in the number of ERASMUS places are a good predictor of international migration (Parey and Waldinger, 2011). Using the selection-corrected Mincer regression we confirm our main results. We also show that our results are not driven by our particular measure of earnings inequality or by potentially confounding factors that may be correlated with cross-country inequality.

Additionally, we show that our results hold for migrants to European countries only. Migration costs to these countries are low because workers can move freely between European countries without the need for work visas. In further results we show that migrants to Austria and Switzerland, two countries with higher earnings inequality than Germany, are positively selected, as predicted by the Roy/Borjas model. These countries provide a useful setting to test for migrant selection because migration costs are particularly low: the two countries share a border with Germany, are predominately German-speaking, and they have broadly similar labor market institutions, benefit systems, and cultures. Furthermore, Germans do not need visas to work in Austria or Switzerland.

In additional results, we decompose predicted earnings to identify the characteristics that explain the observed selection patterns. Migrants to less equal countries have better university grades, attend better universities, and come from families with higher socioeconomic backgrounds. Migrants to more equal countries have studied subjects with lower returns in the labor market, they are more likely to be female, and they attend universities associated with lower labor market prospects. Interestingly, migrants to more equal destinations are, in fact, positively selected in terms of university grade. Selection patterns are thus consistent with the model predictions for most, but not all, characteristics.[5]

[5]A multidimensional extension of the Roy/Borjas model indicates that focusing on a

Predicted earnings provide a comprehensive summary measure of expected productivity that drives migration decisions.

Finally, we investigate selection to the United States, one of the most important destinations of high-skilled emigrants from Germany. In the United States, earnings inequality among university graduates is much higher than in Germany. As predicted by the Roy/Borjas model, emigrants from Germany to the United States are positively selected, compared to non-migrants. We show that migrants to the United States are positively selected across almost all characteristics not only compared to non-migrants in Germany, but also compared to U.S. natives. We also document that migrants from Germany to the United States are particularly concentrated in high-paying STEM fields.

Overall, high-skilled individuals form an important group of potential migrants, both because of their relatively high rates of mobility and their potential contribution to the host economy. Studying migrant selection in this context is particularly useful because these migrants face low formal barriers to migration and because they are unlikely to be credit constrained. The observed selection patterns underline the relevance of the Roy/Borjas model in this setting.

1 A Model of Migrant Selection and Existing Empirical Evidence

1.1 Roy/Borjas Model of Migrant Selection

In his seminal work, Borjas (1987, 1991) proposes a theoretical framework for understanding the selection of international migrants. To motivate our empirical analysis, we use important insights of the Roy/Borjas model to highlight the predictions for selection. In our context, university graduates decide whether to migrate based on earnings opportunities abroad (w_1) and at home (w_0), and migration costs (c). In this framework,

single characteristic may not reflect the overall pattern of selection, depending on the correlation with other relevant characteristics. See Dustmann et al. (2011) for a model with two types of skills.

potential log earnings consist of an observed component (θ_j, where $j = 0$ indicates home and $j = 1$ indicates abroad) and an unobserved component (ϵ_j):

$$\log w_0 = \theta_0 + \epsilon_0 \tag{1}$$

$$\log w_1 = \theta_1 + \epsilon_1. \tag{2}$$

Taking migration costs (c) into account, individuals will move abroad if the wage gain is larger than the migration costs:

$$Migrate{=}1 \text{ if } \theta_1 + \epsilon_1 > \theta_0 + \epsilon_0 + c. \tag{3}$$

The vector of potential outcomes is $(\theta_0, \theta_1, \epsilon_0, \epsilon_1)$. For tractability, we assume that the outcome vector is jointly normally distributed with means $(\mu_0, \mu_1, 0, 0)$ and variances $(\sigma^2_{\theta_0}, \sigma^2_{\theta_1}, \sigma^2_{\epsilon_0}, \sigma^2_{\epsilon_1})$. Mean earnings at home and abroad are represented by μ_j, and the variance of the observed component in each country is represented by $\sigma^2_{\theta_j}$. We allow each type of skills (observables and unobservables) to be correlated across countries, but not across types of skills. $\sigma_{\theta_0,\theta_1}$ is the covariance in the observed component across countries. We refer to the corresponding correlation as ρ_θ. While our framework incorporates observed and unobserved skills, this does not affect the underlying economic mechanism developed by Borjas (1987, 1991).[6]

We now consider how earnings potential at home, θ_0, of migrants differs from the population mean μ_0. From the normality assumption we obtain

$$E(\theta_0|Migrate{=}1) = E(\theta_0|\theta_1 + \epsilon_1 > \theta_0 + \epsilon_0 + c) \tag{4}$$

$$= \mu_0 + \left(\rho_\theta - \frac{\sigma_{\theta_0}}{\sigma_{\theta_1}}\right) \frac{\sigma_{\theta_0}\sigma_{\theta_1}}{\sigma_v} \frac{\phi(z)}{1 - \Phi(z)}, \tag{5}$$

where $v = \theta_1 + \epsilon_1 - \theta_0 - \epsilon_0$ is the earnings difference between abroad and home that has

[6]Borjas (1987) develops the original model focusing on the role of unobservables. In the formulation here, this corresponds to the case of $\sigma_{\theta_0} = \sigma_{\theta_1} = 0$. Borjas (1991) introduces the distinction between returns to observables and unobservables, focusing on the case where observable skills are perfectly correlated across countries ($corr(\theta_0, \theta_1) = 1$).

variance σ_v^2. $z = (\mu_0 + c - \mu_1)/\sigma_v$ is a constant reflecting differences in means across destinations, adjusted for migration costs and normalized by the variance of the earnings difference. In our empirical analysis, we investigate how selection on observables relates to relative inequality ($\sigma_{\theta_0}/\sigma_{\theta_1}$) between the two destinations.[7] In addition to relative inequality, the theoretical prediction on selection depends on the cross-country correlation in the observed component (ρ_θ). A situation where ρ_θ is sufficiently high provides a natural benchmark case because we analyze migration flows between industrialized countries.[8] If the potential destination is *less equal* than home ($\sigma_{\theta_1} > \sigma_{\theta_0}$), migrants will be positively selected: $E(\theta_0|\text{Migrate=1}) > \mu_0$. Intuitively, the positively selected migrants benefit from the upside opportunities in less equal countries. If the potential destination country is *more equal* ($\sigma_{\theta_1} < \sigma_{\theta_0}$), migrants will be negatively selected: $E(\theta_0|\text{Migrate=1}) < \mu_0$. Intuitively, the negatively selected migrants benefit from the insurance of a compressed wage distribution in more equal countries.

The model emphasizes the role of earnings inequality for the selection of migrants. Differences in mean earnings between home and abroad have strong effects on migration probabilities (see term z above), but they have no effect on the *direction* of selection.

Borjas (1991) extends the model to include stochastic migration costs, leading to very similar results as long as the migration costs are unrelated to potential earnings; Chiquiar and Hanson (2005) emphasize that selection patterns can change substantially when migration costs vary systematically with earnings potential. Because we are focusing on high-skilled individuals who migrate from an economically developed country to other developed countries, differential migration costs are presumably less important than for

[7]Our data include a rich set of observable characteristics, which allow us to construct an informative measure of skills. See Gould and Moav (2016) for an analysis that investigates selection on unobservable skills.

[8]This rules out the case of 'refugee sorting' (Borjas 1987).

lower-skilled migrants who, e.g., migrate from Mexico to the United States.

1.2 Empirical Evidence on the Roy/Borjas Model

Most empirical papers on international migrant selection study settings where migrants face legal barriers to migration and migration costs are relatively high. Existing papers differ along two main dimensions that affect observed selection patterns. First, different papers use different skill measures to evaluate selection, and second, they study migration flows between a varying set of countries (see Online Appendix Table A.1). A large part of the empirical literature has studied emigration from Mexico to the United States. While some of these papers find evidence for negative selection that is consistent with the basic Roy/Borjas model (e.g. Ibarraran and Lubotsky, 2007, Fernández-Huertas Moraga, 2011, Kaestner and Malamud, 2014, for some characteristics), other papers find intermediate selection that suggests that migration costs vary with skills, perhaps driven by poverty constraints (Chiquiar and Hanson, 2005, Orrenius and Zavodny, 2005, Kaestner and Malamud, 2014, for other characteristics). In their seminal paper, Chiquiar and Hanson (2005) show that a model with skill-varying migration costs provides a better description of migration flows from Mexico to the United States.

A number of other papers investigate migrant selection between other pairs of countries. The selection of migrants from Puerto Rico to the United States is consistent with the model predictions (Ramos, 1992, Borjas, 2008). Migrant selection from either Norway or Israel to the United States is only partly consistent with the model predictions (Abramitzky et al., 2012, Gould and Moav, 2016).

Lastly, a number of papers investigate migrant selection between multiple countries. Some papers find support for the model predictions (e.g. Borjas, 1987, Borjas et al., 2015, Stolz and Baten, 2012), while other cross-country studies find only partial support for the basic Roy/Borjas model (Belot and Hatton, 2012), or reject the model predictions

(Feliciano, 2005, Grogger and Hanson, 2011).[9]

We are not aware of other papers that focus on the role of inequality for the selection of *high-skilled* migrants.[10] Studying these migrants is particularly useful because they face low legal barriers to migration and relatively small migration costs.

2 Data

2.1 Data on University Graduates

We analyze the selection of high-skilled migrants using survey data on university graduates collected by the German Centre for Higher Education Research and Science Studies (DZHW). These data come from nationally representative longitudinal surveys of individuals who complete their university education in Germany (for details see Grotheer et al., 2012). The DZHW sampled university graduates from the graduation cohorts 1992-93, 1996-97, 2000-01, and 2004-05.[11] We refer to the cohorts by the second year, i.e. 1993 for the 1992-93 cohort. Graduates in each cohort are surveyed twice. The initial survey takes place about 12 months after graduation. The same individuals participate in a follow-up survey about five years after graduation (Online Appendix Figure A.1). The survey is ideal for our purposes because graduates are surveyed even if they move abroad. We focus our analysis on migration decisions that are measured five years after

[9]Our focus is on the selection of *international* migrants. A number of papers investigate the Roy/Borjas model applied to internal migration, including Borjas et al. (1992), Dahl (2002), Abramitzky (2009), and Bartolucci et al. (2014).

[10]Recent papers have highlighted the role of taxes for the migration of inventors and soccer players (Akcigit et al., 2016, Kleven et al., 2013).

[11]Between 1993 and 2005, the majority of German university graduates completed degrees called *Diplom*, *Magister*, or *Staatsexamen*. These degrees are usually completed in four to six years, and are considered comparable to a master's degree in other countries in standard international classifications (ISCED 5A according to the International Standard Classification of Education).

graduation.[12] The surveys are based on a stratified cluster sampling, with fields of study, degree types, and universities as strata (Grotheer et al., 2012), and they are representative for the examined population. Response rates to the initial surveys range between 30% and 40%, depending on the cohort. We analyze differences in response rates between the initial survey and the follow-up survey according to migration status reported in the initial survey. The follow-up survey response rate is 66% for graduates who have worked in Germany during the initial survey and 59% for graduates who have worked abroad. While this difference is statistically significant in a simple t-test, we cannot reject that differences in response rates are uncorrelated to observable characteristics. This suggests that attrition does not change our findings. We also verify that our results hold when we include the full sample from the initial survey by carrying forward the reponses from the initial survey (see Online Appendix A.3 for details, results are shown in Online Appendix Figure A.3 and in Online Appendix Table A.6).

Five years after graduation, the total number of respondents is 6,737 (1993 cohort), 6,237 (1997 cohort), 5,426 (2001 cohort), and 6,459 (2005 cohort). We focus on graduates from traditional universities.[13] We restrict the sample to full-time workers because

[12]After graduation, many university graduates enroll in additional training such as legal or teacher traineeships (*Referendariat*), or PhD programs. Earnings in the initial survey are thus a noisy measure of earnings potential.

[13]The German higher education sector consists of traditional universities, universities of applied sciences (*Fachhochschulen*), specialized universities (focusing on arts, music, or theology), and a small number of private universities. The best students tend to enroll in traditional universities. To estimate the Heckman selection model we also restrict the sample to graduates from universities where at least one graduate works abroad. These sample restrictions reduce the sample by around 30%. Results that include all institutions are very similar to our main findings (see Online Appendix Figure A.4).

migrating part-time workers are more likely to be tied movers (see Borjas and Bronars, 1991; Junge et al., 2014). In our data, full-time labor force participation is about 77%.

The graduate survey data contain detailed information on graduates' personal characteristics, family background, study history, and labor market experience (Table 1). In addition to the variables summarized in Table 1, we also have detailed information on a graduate's university and field of study. Five years after graduation, 5.2% of graduates work abroad. The main destinations are Switzerland, the United States, the UK, Austria, and France (Online Appendix Table A.2).

2.2 Data on Earnings Inequality

We classify destination countries as either more or less equal than Germany using newly constructed measures of earnings inequality for university graduates. Existing inequality measures, such as Gini coefficients, typically measure inequality for the overall population, but the decisions of high-skilled migrants will likely depend on earnings inequality of university graduates.

Our main data source is the Luxembourg Income Study (LIS) (2013). The LIS provides access to individual-level earnings surveys from several countries. The database covers different years for each country. We use all available survey years for the main destinations of German university graduates (see Table A.3 for available survey years in each country). Switzerland and Austria are important destinations for German university graduates but only have relatively limited coverage in the LIS database. We therefore augment the LIS data with additional data for Austria (Microcensus 1999 and EU-SILC 2007, 2008) and Switzerland (Labour Force Survey 1998-2005).

To measure earnings inequality for high-skilled individuals, we restrict the samples in the individual-level income surveys to university graduates. We further restrict the samples to full-time employees of working age, and we exclude individuals who are self-employed, enrolled in educational institutions, or who report negative earnings.

Based on the individual-level surveys, we construct earnings percentiles for each country and available year using the survey sampling weights (see Online Appendix Table A.3). Some surveys in the (augmented) LIS data report gross earnings, while others report net earnings. To measure cross-country inequality of net earnings, we convert gross into net earnings using the *net personal* average tax rate of single persons without children from the OECD (2013c).[14] The Data Appendix B.1 provides more detail on the construction of the inequality measures.

In our main analysis, we use the ratio of the 75th to the 25th earnings percentile (75/25 ratio) for university graduates to measure earnings inequality across countries. Figure 1 shows the ranking of countries according to the 75/25 ratio that we average over 1998 to 2010 to reflect the years that correspond to our graduate surveys (Online Appendix Table A.4 reports 75/25 ratios for each country). Inequality is highest in the United States, followed by France and Poland. The Scandinavian countries and Australia are most equal. Germany is ranked in the middle.[15] We can therefore investigate the

[14]The net personal average tax rate is defined as the personal income tax and employee social security contributions net of cash benefits, expressed as a percentage of gross wage earnings. The OECD reports three different tax rates along the earnings distribution: the average tax rate at 67%, at 100%, and at 167% of average earnings. We apply the tax rate at 67% of average earnings to the 25th percentile and below, the tax rate at 100% of average earnings to earnings between the 25th and the 75th percentile, and the tax rate at 167% of average earnings to the 75th percentile and above.

[15]Recent papers have used large administrative datasets to documente a rise in German earnings inequality during the last decades (Dustmann et al., 2009, Card et al., 2013). In these datasets, earnings are censored at the maximum of social security contributions. For university graduates, 42% (13%) of observations for males (females) are top-coded between 1998 and 2008. Because university graduates are in the top 11% of the educational

selection of German university graduates into less equal and into more equal countries.[16]

2.3 Data on ERASMUS Places

As part of our estimation procedure, which we explain below, we use data on the number of ERASMUS places to correct for potential selection bias. We obtain data on the number of study abroad places in the ERASMUS program by university, subject, and year from the German Academic Exchange Service (DAAD). The median internationally mobile student studies abroad for one or two semesters about three years before graduation. We assign the number of ERASMUS places in the corresponding academic year, subject, and university to each student. To account for differences in cohort size that affect students' study abroad opportunities, we normalize the number of ERASMUS places with the number of students in the corresponding university and subject (for details see Parey and Waldinger, 2011).

3 Methods and Results

3.1 The Selection of Migrants to More and to Less Equal Destinations

For our analysis, we use predicted earnings to measure earnings potential in the home country. This measure of skill represents θ_0 in the model outlined above. We then use predicted earnings to compare the distribution of skills of migrants to less equal countries, of migrants to more equal countries, and of non-migrants.

To construct predicted earnings, we estimate an augmented Mincer regression for

distribution, we prefer to use earnings surveys in the LIS that are not top-coded.

[16] As we measure selection with predicted earnings, an ideal measure of inequality would be based on country-level differences in returns to observed skills. Such a measure would require graduate datasets with comparable characteristics on each graduate for all major destinations. As these are not available, we use the 75/25 ratio that is based on actual earnings. The empirical results are valid as long as countries with higher 75/25 ratios also exhibit higher returns to observed skills.

non-migrants only:

$$\log w_{0i} = X_i\beta_0 + \varepsilon_{0i} \tag{6}$$

The estimate of β_0 measures returns to skills in the home country. Our data allow us to include a large number of variables X_i to obtain a good prediction of earnings potential. X_i contains variables that measure university experience (final university grade, age at graduation, completing university with a bachelor's degree, 24 subject fixed effects, and university fixed effects), additional education after graduation (completing a PhD or a non-PhD graduate degree), pre-university education (final high school grade and apprenticeship before studying), previous mobility (moving to another state between high school and university), potential labor market experience, personal characteristics (gender, marital/partnership status, children), parental background (mother's and father's education and occupation), and graduate cohort fixed effects. The coefficients of the augmented Mincer regression have the expected signs and magnitudes (Table 2, column 1).[17] The R^2 of about 0.28 is high for a Mincer regression, suggesting that predicted earnings are an informative skill measure for university graduates.[18]

Next, we predict potential earnings in the home country for migrants and non-migrants. The predictions are based on the coefficient vector ($\hat{\beta}_0$) and on individual characteristics X_i.

$$\hat{\theta}_{0i} = X_i\hat{\beta}_0 \tag{7}$$

[17]Because all graduates are surveyed around five years after graduation, the variation in potential labor market experience is small and estimated coefficients are different from the typical pattern observed in Mincer regressions. The omitted degree is a Diplom/Magister degree. Compared to graduates with these traditional degrees, graduates sampled after completing a Bachelor's degree have lower earnings.

[18]We show that our results are robust to excluding the controls for children and marital/partnership status from the wage regression (Online Appendix Figure A.5).

We then use this measure of skills to compare three groups of interest: migrants to less equal countries, migrants to more equal countries, and non-migrants. Specifically, we construct Cumulative Distribution Functions (CDFs) of predicted earnings $\hat{\theta}_0$ by migration group, $F(\hat{\theta}_0 \,|\, \text{Migration status})$, and plot them in the left panel of Figure 2(a).

The dashed line shows the CDF of non-migrants. The dark, solid line is the CDF of migrants to *less equal* destinations, such as the United States. This CDF lies to the right of the CDF for non-migrants, indicating that this group is *positively* selected in terms of earnings potential. The migrants to these countries have skills which, according to the returns in the Mincer regression, are valued more highly than those of non-migrants: median log predicted earnings for these migrants are 10.65, compared to 10.61 for non-migrants. The CDFs of non-migrants and of migrants to less equal countries do not cross, indicating that these migrants are positively selected over the full range of predicted earnings. We test the statistical significance of our findings in section 3.3.

The lighter, solid line shows the CDF of migrants to *more equal* destinations, such as Sweden. It indicates that migrants to more equal countries are *negatively* selected relative to non-migrants. Median log predicted earnings for these migrants are 10.56, compared to 10.61 for non-migrants. The differences between the CDFs are substantial and in the same order of magnitude as standard estimates for the returns to an additional year of education in the United States (Card, 1999).

Inequality varies across potential destination countries. We use this variation to analyze selection to countries with more extreme levels of (in)equality by splitting more and less equal countries into two groups each. Thus, we now compare five types of destinations: very unequal, somewhat unequal, home, somewhat equal, and very equal countries. We classify the three countries with the most unequal earnings distributions as very unequal, and the three countries with the most equal distributions as very equal. Results are shown in the right panel of Figure 2(a). Very unequal countries receive the most

positively selected migrants; somewhat unequal countries receive somewhat positively selected migrants; somewhat equal countries receive slightly negatively selected migrants; and very equal countries receive strongly negatively selected migrants. The CDFs are somewhat noisier than in the previous graphs because sample sizes of migrants are relatively small, especially for equal countries. Nonetheless, the selection pattern follows the theoretical predictions for the five groups.[19]

3.2 Controlling for Selection in the Augmented Mincer Regression

As our previous analysis has shown, migrants are systematically selected from the home population. Unless this selection is fully accounted for by the observables, the selection could potentially bias the coefficients of the augmented Mincer regression and thus our measure of predicted earnings. We use a Heckman selection procedure to control for this potential selection by estimating a selection equation that predicts whether a graduate works in Germany or migrates abroad.

We use the introduction and expansion of the ERASMUS student exchange program as an instrumental variable that predicts whether graduates work in Germany. ERASMUS allows students to study abroad in a European country for one or two semesters before they continue their studies in their home country. It was introduced in 1987 and increased massively since then. In Germany, about 4,925 students participated in ERASMUS in 1990 (the year when the typical graduate of the 1993 cohort had studied abroad), and participation rose to 18,482 in 2002 (the year when the typical graduate of the 2005 cohort had studied abroad). The program was introduced at different times and expanded at varying rates, depending on the university and department. Parey and Waldinger (2011) show that the introduction and expansion of ERASMUS significantly increased the probability of graduates moving abroad after completing their studies in Germany.

[19] In Online Appendix Figure A.6, we show results where we classify the four, instead of three, most (un)equal countries as most (un)equal. The results are very similar.

The ERASMUS instrument successfully controls for selection in the Mincer regression if the number of ERASMUS places in a student's university can be excluded from the Mincer regression. Crucially, we do not use the actual decision to study abroad, but the availability of department-level ERASMUS scholarship places, which predict studying abroad and working abroad later on, to instrument for working in Germany.

In our data, the median graduate enrolled in university in 1991-92 and thus before the widespread availability of the Internet. Before the introduction of the Internet, information on the number of ERASMUS places was very difficult to obtain. Even today, few department websites report the exact number of ERASMUS places. It is therefore unlikely that students sorted into certain departments to benefit from more ERASMUS places. To further limit the possiblity of student sorting, we assign the number of ERASMUS places for the subject×university combination where a student first enrolls in university. Any potential sorting after the first enrollment will therefore not affect the exogeneity of the ERASMUS instrument.

Students in certain subjects are systematically more likely to study abroad, and to work abroad later on, than students in other subjects. We control for any such subject-specific differences by including 24 subject fixed effects in the regressions. A related concern may be that better universities offer more ERASMUS places and also facilitate working abroad. We control for these university-specific differences by including a full set of university fixed effects in the regressions. We also control for broader trends of studying and working abroad by controlling for cohort fixed effects.

Parey and Waldinger (2011) further discuss the exclusion restriction of the ERASMUS instrument. They show that the expansion of ERASMUS in a department is not correlated with a wider push to increase the international outlook of students. They also show that the probability of studying abroad is completely flat before the introduction of ERASMUS and only increases once ERASMUS has been introduced, suggesting that

pre-trends are not affecting the validity of the ERASMUS instrument.

Column (3) of Table 2 shows the first-stage estimates where we regress whether individuals work in Germany on a measure of ERASMUS scholarship places (normalized by the number of students) in a graduate's university department. The availability of ERASMUS significantly lowers the probability of working in Germany.

Column (2) in Table 2 shows that controlling for selection in the Mincer regression only has a small effect on the estimated coefficients. In addition to the rich set of observables, this also reflects that the share of graduates not migrating (and thus observed in our Mincer regression) is very high, and that selection of migrants occurs both at the top and the bottom of the distribution. The coefficient on the Mills ratio is therefore quantitatively small and insignificant. The resulting CDFs of earnings potential by migration status are presented in Figure 2(b). They confirm that migrants to less equal destinations are positively selected, while migrants to more equal destinations are negatively selected.

3.3 Tests for Stochastic Dominance

We investigate the statistical significance of the substantial differences between the CDFs with tests for first-order stochastic dominance. As we estimate the Mincer earnings equation in the first step of our analysis and construct predicted earnings based on the Mincer regression, we need to account for this additional source of uncertainty when we compute p-values. We therefore apply the bootstrap procedure for stochastic dominance tests developed in Barrett and Donald (2003) and described in further detail in Online Appendix A.2. We also report p-values from conventional Kolmogorov-Smirnov tests, which do not account for the uncertainty associated with the estimation of parameters in the Mincer regression.

The corresponding test results are shown in Table 3. The top row of panel A indicates that we can reject the null hypothesis that the more-equal-CDF dominates the CDF of non-migrants ('Home') at the 1% level of significance (columns (1) to (3)). Similarly, the

second row indicates that we can reject that the CDF of non-migrants dominates the more-unequal-CDF at the 10% level. We also reject that the more-equal-CDF dominates the less-equal-CDF at the 1% level. We even reject these hypotheses when we use the Heckman selection-corrected estimates, as reported in columns (4) to (6).

The graphical analysis presented above suggests even more pronounced differences in the CDFs when we limit the comparison to very equal and very unequal countries, respectively. Panel B of Table 3 indeed shows that the test statistic for the comparison of these more extreme destinations increases substantially. Because the relevant samples become smaller for destinations with more extreme levels of inequality, the p-values do not decrease in all cases. Nonetheless, the test of stochastic dominance now rejects at the 5% level for all three comparisons. Panel B of Table 3 also reports tests for selection between more similar destinations. The test statistic always has the predicted sign, suggesting that selection follows the basic Roy/Borjas model, even for more similar destinations. As expected, selection patterns to the more similar destinations are often not statistically significant because inequality differences in more similar destinations are much lower and because some country groups attract relatively few graduates. We also test the reverse set of hypotheses and cannot reject them. The corresponding p-values are above 0.74 and in most cases above 0.95 (Online Appendix Table A.5).

3.4 Selection of Migrants by Country

Our data also allow us to investigate the selection of migrants to each of the 19 destinations in our sample, and thereby go beyond the three or five groups of countries presented in the previous section. We compute average predicted earnings of migrants to each country and correlate them with the 75/25 ratio (Figure 3). Circle sizes indicate the number of migrants in each country. Apart from a few outliers, migrants to more equal countries have lower predicted earnings than migrants to less equal countries.

We estimate a weighted country-level OLS regression and show the corresponding

prediction in Figure 3. In particular, we regress average predicted earnings ($\bar{\hat{\theta}}_{0c}$) on the 75/25 *ratio* in each country c:

$$\bar{\hat{\theta}}_{0c} = \gamma_0 + \gamma_1 75/25\ ratio_c + \varepsilon_c \tag{8}$$

The estimated regression line ($\hat{\gamma}_1$) has a slope of 0.153 with a standard error of 0.081 (Table 4, column (1), significant at the 10% level). This estimate indicates that migrants to destinations with a 75/25 ratio that is higher by 0.4 (the difference between Germany and the United States) have predicted earnings that are 6.1 log points higher.

4 Robustness

4.1 Controlling for Possible Confounding Factors

The selection pattern described in the previous section is consistent with the theoretical predictions of the Roy/Borjas model. Earnings inequality, however, is not the only factor that differs between home and destination countries. Countries may also differ along other dimensions that could be correlated with migrant selection.

We first analyze whether confounding factors (F_c) are driving our selection results by controlling for them in the cross-country regression (Table 4):

$$\bar{\hat{\theta}}_{0c} = \gamma_0 + \gamma_1 75/25\ ratio_c + \gamma_2\, F_c + \varepsilon_c \tag{9}$$

The Roy/Borjas model predicts that *mean* earnings should affect the number of migrants to each country but not the direction of selection. Nonetheless, differences in mean earnings will affect migration choices and may be correlated with differences in the 75/25 ratios. In our first robustness check, we therefore control for average log earnings in each country. In this specification, the coefficient on the 75/25 ratio increases slightly to 0.180, suggesting an even stronger relationship between inequality and migrant selection (column (2), significant at the 1% level). Migration decisions, especially those of lower-skilled migrants (within the high-skilled population), may also be affected by expected unemployment spells. When we control for unemployment rates of tertiary-educated individuals the coefficient on the 75/25 ratio is equal to 0.174 (column (3), significant at the

5% level). Migration decisions may also be affected by differences in child-care provision. When we control for public expenditures on family benefits the coefficient on the 75/25 ratio is equal to 0.110 (column (4), significant at the 10% level). Migration decisions may also be affected by expectations about general well-being. When we control for a measure of life satisfaction the coefficient on the 75/25 ratio is 0.247, confirming a strong relationship between earnings inequality and migrant selection (column (5), significant at the 1% level). When we control for all potential confounders at the same time the coefficient on the 75/25 ratio is 0.147, with a p-value of 0.061 (column (6)).

The previous checks confirm a robust effect of earnings inequality on *mean* selection levels. In additional tests, we investigate how potential confounders affect selection across the whole distribution of skills. For these tests, we first replicate the CDFs from our main results using quantile regressions, and then control for possible confounding factors using the quantile regression framework. We regress predicted earnings of each individual i ($\hat{\theta}_{0i}$) on country group dummies separately for 100 centiles ($\tau = 0.01...0.99$) of the predicted earnings distribution:

$$\begin{aligned}\hat{\theta}_{0ic} =& \delta_{0\tau} + \delta_{1\tau} \textit{Very Equal}_{ic} + \delta_{2\tau} \textit{Somewhat Equal}_{ic} + \\ & \delta_{3\tau} \textit{Somewhat Unequal}_{ic} + \delta_{4\tau} \textit{Very Unequal}_{ic} + \epsilon_{ic\tau}\end{aligned} \tag{10}$$

$\textit{Very Equal}_{ic}$ takes a value of 1 if the individual works in a country that is much more equal than Germany, $\textit{Somewhat Equal}_{ic}$ takes a value of 1 if the individual works in a country that is somewhat more equal, and so on. The constant represents predicted earnings for individuals who work in Germany. Online Appendix Figure A.7(a) shows the quantile regression equivalents of the CDFs in our main results. We then control for potential confounding factors in the quantile regressions by adding country-level controls:

$$\begin{aligned}\hat{\theta}_{0ic} =& \delta_{0\tau} + \delta_{1\tau} \textit{Very Equal}_{ic} + \delta_{2\tau} \textit{Somewhat Equal}_{ic} + \\ & \delta_{3\tau} \textit{Somewhat Unequal}_{ic} + \delta_{4\tau} \textit{Very Unequal}_{ic} + \delta_{5\tau} F_c + \epsilon_{ic\tau}\end{aligned} \tag{11}$$

From the estimated coefficients, we predict CDFs for each group holding constant the

value of the added covariate at the German level. Panels (b) to (f) of Online Appendix Figure A.7 show CDFs that are adjusted for the same confounding factors that we have analyzed in the cross-country regression (Table 4). The selection pattern to locations with more extreme levels of (in)equality is robust to controlling for potentially confounding factors. The selection pattern to locations with less extreme levels of (in)equality remains broadly consistent with the predictions of the model (see Online Appendix Table A.7 for stochastic dominance tests). If we control for mean earnings, the CDF for *somewhat unequal* countries sometimes moves to the left of the CDF for graduates at *home*. However, the stochastic dominance tests indicate that the two CDFs are not significantly different.

4.2 Sensitivity of Results to Alternative Inequality Measures

In Appendix Section A.4, we investigate the sensitivity of our main results to using alternative measures of inequality. We show that the results are very similar for a range of inequality measures, including the 90/50 ratio, the 75/25 ratio, the 90/10 inequality ratio, the Gini coefficient and the Theil index.

4.3 Selection to Europe and to Austria/Switzerland

Additionally, we investigate selection to European countries only. German citizens who migrate to these countries face virtually no migration barriers, such as visa requirements. Germans can settle freely in any country of the European Union and in other European countries, such as Switzerland, Liechtenstein, and Norway.[20] Furthermore, migration costs to these countries are relatively low because distances within Europe are small, and travel costs are low.

We plot CDFs of predicted earnings of migrants to less equal countries, migrants to more equal countries, and non-migrants (Figure 4(a)). As for the full sample, migrants to more equal European countries are negatively selected, and migrants to less equal

[20]Graduates from the 1993 to 2001 cohorts who migrated to Poland or Switzerland had minor restrictions to settle in these countries.

European destinations are positively selected, compared to non-migrants.[21] These results suggest that differential migration costs are not driving our main results.

We also investigate migrant selection to Austria and Switzerland only. These two countries are very similar to Germany along many dimensions that may affect migration choices. The countries have similar education systems with very similar university graduation rates (OECD, 2013b, p. 61). The countries also have similar unemployment benefits as measured by replacement rates that ranged between 29% and 33% of gross incomes in 2005 (OECD, 2015). The three countries also share a similar culture. Finally, Austria is German speaking and in Switzerland 64% of the population is German-speaking, and more than 90% of Germans migrants settle in predominately German-speaking regions of Switzerland (BFS, 2010, 2013). While the three countries are similar along many dimensions, they differ in earnings inequality. Both Austria and Switzerland are less equal than Germany. The CDF of predicted earnings of migrants to Austria and Switzerland lies to the right of the non-migrant CDF (Figure 4(b)).[22] These results indicate that migrants to Austria and Switzerland are positively selected compared to non-migrants, as predicted by the Roy/Borjas model.

5 Further Results

5.1 Decomposing Migrant Selection

Predicted earnings can be considered a summary measure of different skills. To understand the characteristics that underlie the observed selection, we use a Blinder-Oaxaca

[21] We reject that the CDF of migrants to more equal countries dominates the home CDF at the 5% level (Online Appendix Table A.9). As Europe contains few countries with very high inequality, we no longer reject that the home CDF dominates the CDF of migrants to less equal countries (p-value of 0.19).

[22] The test that the home CDF dominates the Austria/Switzerland CDF is rejected at the 10% level (Online Appendix Table A.9, panel B).

procedure, decomposing the overall difference in predicted earnings into the contibution of each characteristic. For expositional purposes we group characteristics into 13 categories.

The positive selection of migrants to *less equal* countries mostly reflects their university career (Figure 5, panel (a1)). They have better grades and attend better universities than non-migrants. The negative selection of migrants to *more equal* countries reflects their university subject, university quality, and gender (panel (a2)). They study subjects with lower returns in the labor market, enroll at universities with less favorable labor market prospects, and are more often female. Interestingly, migrants to more equal countries have better grades at university, despite being negatively selected overall. This is consistent with findings suggesting that migrants are positively selected when skill is measured in terms of education. Decomposition results that use coefficients from the selection-corrected Mincer regression are shown in Online Appendix Figure A.9.

Columns (1) to (4) of Table 5 summarize how the covariates of the decomposition line up with the overall prediction. For most characteristics, the table shows no significant deviations from the model predictions. However, there are a number of interesting differences between the relevance of individual characteristics between less equal and more equal countries. For less equal countries, the pattern of selection in terms of apprenticeship training is not in line with our baseline prediction, and for more equal countries, university grade shows significant positive selection among the migrants. Although we do not have the detailed data to investigate these instances, they may reflect heterogeneity in returns to characteristics across countries or a correlation of these characteristics with the willingness to move, in a way not captured by the model. For example, it is plausible that (former) apprentices may realize a higher return to their training in their home labor market and are therefore more attached to their home labor market.

It is important to keep in mind that Figure 5 shows the results of a statistical decomposition and that the characteristics may be correlated with each other. Predicted

earnings provide a natural way of combining the individual characteristics in a summary measure.

5.2 Migration to the United States

Migrants to the United States Compared to Non-Migrants in Germany

In the final section, we investigate migrant selection to the United States. The United States is an important destination for university graduates from Germany. In our sample, more than 13% of graduates who go abroad move to the United States; only Switzerland attracts more graduates from Germany. Because U.S. inequality is highest among the major destinations of German university graduates, we expect that German university graduates who migrate to the United States are particularly positively selected.

The CDF of migrants to the United States always lies to the right of the non-migrant CDF (left panel of Figure 6(a)). The difference between the CDFs of U.S. migrants and non-migrants is more pronounced than the difference between the CDFs of all migrants to less equal countries and non-migrants. This highlights the particularly positive selection of migrants to the United States. The test of stochastic dominance is rejected at the 5% level (see Online Appendix Table A.9, panel C).

As above, we decompose the difference in predicted earnings between migrants to the United States and non-migrants. U.S. migrants are positively selected according to almost all characteristics, in particular characteristics that relate to the university career and gender. Migrants to the United States study subjects with especially high returns (see third bar from the top in the right panel of Figure 6(a)). They are particularly concentrated in STEM fields. In our sample, about 17.2% of migrants to the United States hold a degree in physics (but only 3.9% of non-migrants), 9.2% hold a degree in biology (non-migrants: 2.3%), and 8.1% hold a degree in chemistry (non-migrants: 3.0%). Furthermore, migrants to the United States are also more likely to hold degrees in computer science, economics and management, geography, and engineering; and they

are less likely to hold degrees in law, languages, medicine, architecture, and education. Migrants to the United States also obtain higher grades in university than non-migrants. They also study in universities where graduates have higher predicted earnings.

The decomposition indicates that the United States attracts high-skilled migrants from Germany who have studied in better universities, received higher grades, and are concentrated in high-paying STEM fields. Thus, migrants to the United States are precisely those that are considered to be important for innovation and technological progress.

Migrants from Germany Compared to U.S. Natives in the ACS

Finally, we investigate how high-skilled migrants from Germany fare in the U.S. labor market by comparing earnings potential of high-skilled migrants from Germany to high-skilled *natives* in the United States. For this test, we use data from the American Community Survey (ACS), and identify high-skilled migrants from Germany as individuals who were born in Germany to non-U.S. parents, who migrated to the United States between 1996 and 2010 and were at least 25 years old at the time of migration. These restrictions ensure that our sample of Germans in the United States is as similar as possible to the sample of graduate emigrants from Germany who we study in our main results. To focus our analysis on the high-skilled, we limit the sample to individuals with a bachelor's degree or higher, who worked for 50 to 52 weeks per year in full-time jobs, and who are 30 to 45 years old (see Data Appendix B.1.4 for further details on the ACS data).[23]

We then compare predicted earnings of migrants from Germany to earnings of U.S. natives. We evaluate the skills of German immigrants to the United States using predicted earnings that we construct from returns to skills for U.S. natives (see Online Appendix Table A.10, column (1) for returns to skills for U.S. natives). In terms of the Roy/Borjas

[23]Results are similar if we restrict the ACS sample to graduates in more academically oriented subjects to further increase the comparability with graduates from traditional universities in Germany.

model presented above, this test compares the distribution of $\hat{\theta}_1$ of German migrants in the United States to U.S. natives, while our previous results compared distributions of $\hat{\theta}_0$ of migrants and non-migrants.[24] Indeed, our results show that compared to high-skilled U.S. natives, recent migrants from Germany have far higher predicted earnings in the U.S. labor market. The CDF of predicted earnings of German immigrants lies to the right of the native CDF along the whole earnings distribution (left panel of Figure 6(b)). At the median, log predicted earnings of migrants from Germany are 11.383, while log earnings of natives are 11.129. At the 25th and 75th percentiles, migrants from Germany have predicted earnings of 11.193 and 11.554, while natives have predicted earnings of 10.937 and 11.334. A back-of-the-envelope calculation suggests that the stronger degree selection in terms of θ_1 (relative to our earlier results in terms of θ_0) can be reconciled with our theoretical prediction, both qualitatively and quantitatively.[25]

Overall, these results indicate that high-skilled individuals who migrate from Germany to the United States are not only positively selected compared to Germans who do not migrate, but also compared to non-migrants in the United States. To investigate the contribution of different characteristics, we also decompose the difference in predicted earnings between German migrants to the United States and U.S. natives. Because the

[24]Parallel to equation (5), the corresponding equation for selection in terms of earnings potential in the destination country is $E(\theta_1|Migrate{=}1) = \mu_1 + \left(\frac{\sigma_{\theta_1}}{\sigma_{\theta_0}} - \rho_\theta\right) \frac{\sigma_{\theta_0}\sigma_{\theta_1}}{\sigma_v} \frac{\phi(z)}{1-\Phi(z)}$.

[25]The selection in terms of θ_1 should be stronger than selection in terms of θ_0 by a factor of $\left(\sigma_{\theta_1}/\sigma_{\theta_0} - \rho_\theta\right) / \left(\rho_\theta - \sigma_{\theta_0}/\sigma_{\theta_1}\right)$. Between the United States and Germany, the ratio $\sigma_{\theta_0}/\sigma_{\theta_1}$ is about 0.8 in our data. While ρ_θ is unknown, the positive selection in terms of θ_0 indicates that ρ_θ is above 0.8 (from equation (5)). Suppose $\rho_\theta = 0.9$, then the factor results in a value of 3.7, which is broadly similar but slightly larger than the observed difference in selection. Because the factor decreases in ρ_θ, it is straightforward to reconcile the observed difference in selection with a value of ρ_θ somewhat above 0.9.

ACS data are less detailed than our graduate survey data, the decomposition involves fewer characteristics. Compared to U.S. natives, German migrants have more advanced degrees (such as professional degrees or PhDs) and graduated in subjects (in particular STEM subjects) that typically lead to higher-paid employment. German migrants are also less likely to be female than U.S. natives. Overall, the positive selection compared to U.S. natives reflects similar characteristics as the ones we find for the positive selection compared to German non-migrants.

6 Conclusion

The seminal work of Borjas has emphasized how migrant selection is driven by inequality in home and destination countries: high-skilled individuals benefit from the upside opportunities in less equal countries, and low-skilled individuals benefit from the insurance of a more compressed wage distribution in more equal countries. This insight has motivated various empirical tests of the Borjas model. In spite of the large differences in inequality across many home-destination country pairs, the empirical evidence is mixed.

In this paper, we investigate selection within the group of *high-skilled* migrants in a setting where migration costs are particularly low. We use predicted wages to measure the skills of migrants and graduates who remain at home. Consistent with the predictions of the basic Roy/Borjas model, we find that migrants to more equal countries, such as Denmark, are negatively selected compared to non-migrants. Migrants to less equal countries, such as the United States, are positively selected. In further results we show that migrant selection follows the predictions of the Roy/Borjas model even within subgroups of either more or less equal countries.

Our results are robust to controlling for potentially confounding factors and to using alternative measures of inequality in destination countries. We also demonstrate that the selection pattern holds when we study migration within Europe, and migration to Austria and Switzerland only, where barriers to migration are particularly low. When

we decompose predicted earnings into various skill components, we find that selection patterns follow the model prediction for most, but not all, characteristics, suggesting that the choice of the skill measure can affect findings of migrant selection.

Overall, our findings highlight the importance of the Roy/Borjas model for the selection of high-skilled migrants.

References

Abramitzky, R. (2009). The Effect of Redistribution on Migration: Evidence from the Israeli Kibbutz. *Journal of Public Economics*, 93(3-4):498–511.

Abramitzky, R., Boustan, L. P., and Eriksson, K. (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.

Akcigit, U., Baslandze, S., and Stantcheva, S. (2016). Taxation and the International Mobility of Inventors. *American Economic Review*, forthcoming.

Barrett, G. F. and Donald, S. G. (2003). Consistent Tests for Stochastic Dominance. *Econometrica*, 71(1):71–104.

Bartolucci, C., Villosio, C., and Wagner, M. (2014). Who Migrates and Why? Evidence from Italian Administrative Data. Working Paper.

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

BFS (2010). Ständige und Nichtständige Wohnbevölkerung nach Kanton, Geschlecht, Anwesenheitsbewilligung, Altersklasse und Staatsangehörigkeit. http://www.bfs.admin.ch/bfs/portal/de/index/themen/01/07/blank/data/01.html.

BFS (2013). Ständige Wohnbevölkerung nach Kanton und Hauptsprachen. http://www.bfs.admin.ch/bfs/portal/de/index/themen/01/05/blank/key/sprachen.html.

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

Borjas, G. J. (2008). Labor Outflows and Labor Inflows in Puerto Rico. *Journal of Human Capital*, 2(1):32–68.

Borjas, G. J. and Bronars, S. G. (1991). Immigration and the Family. *Journal of Labor Economics*, 9(2):123–148.

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

Borjas, G. J. and Doran, K. B. (2012). The Collapse of the Soviet Union and the Productivity of American Mathematicians. *Quarterly Journal of Economics*, 127(3):1143–1203.

Borjas, G. J., Kauppinen, I., and Poutvaara, P. (2015). Self-Selection of Emigrants: Theory and Evidence on Stochastic Dominance in Observable and Unobservable Characteristics. NBER Working Paper No. 21649.

Brücker, H., Bertoli, S., Facchini, G., Mayda, A. M., and Peri, G. (2012). Understanding Highly Skilled Migration in Developed Countries: The Upcoming Battle for Brains. In Boeri, T., Brücker, H., Docquier, F., and Rapoport, H., editors, *Brain Drain and Brain Gain - The Global Competition to Attract High-Skilled Migrants*, pages 15–188. Oxford University Press.

Card, D. (1999). The Causal Effect of Education on Earnings. In Ashenfelter, O. C. and Card, D., editors, *Handbook of Labor Economics*, volume 3A, chapter 30, pages 1801–1863. Elsevier, North-Holland.

Card, D., Heining, J., and Kline, P. (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *Quarterly Journal of Economics*, 128(3):967–1015.

Chambers, E. G., Foulon, M., Handfield-Jones, H., Hankin, S. M., and Michaels, E. G.

(1998). The War for Talent. *McKinsey Quarterly*, 3:44–57.

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

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

DESTATIS (2013). Bildungsstand der Bevölkerung. Federal Statistical Office of Germany.

Docquier, F. and Rapoport, H. (2012). Globalization, Brain Drain, and Development. *Journal of Economic Literature*, 50(3):681–730.

Doran, K., Gelber, A., and Isen, A. (2014). The Effects of High-Skilled Immigration on Firms: Evidence from H-1B Visa Lotteries. NBER Working Paper No. 20668.

Dustmann, C., Fadlon, I., and Weiss, Y. (2011). Return Migration, Human Capital Accumulation and the Brain Drain. *Journal of Development Economics*, 95(1):58–67.

Dustmann, C., Ludsteck, J., and Schönberg, U. (2009). Revisiting the German Wage Structure. *Quarterly Journal of Economics*, 124(2):843–881.

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). New Evidence on Emigrant Selection. *Review of Economics and Statistics*, 93(1):72–96.

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

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

Grotheer, M., Isleib, S., Netz, N., and Briedis, K. (2012). *Hochqualifiziert und gefragt. Ergebnisse der zweiten HIS-HF Absolventenbefragung des Jahrgangs 2005*. HIS GmbH.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, 47(1):153–161.

Hunt, J. and Gauthier-Loiselle, M. (2010). How Much Does Immigration Boost Innovation? *American Economic Journal: Macroeconomics*, 2(2):31–56.

Ibarraran, P. and Lubotsky, D. (2007). Mexican Immigration and Self-Selection: New Evidence from the 2000 Mexican Census. In Borjas, G. J., editor, *Mexican Immigration to the United States*, chapter 5, pages 159–192. University of Chicago Press.

Junge, M., Munk, M., and Poutvaara, P. (2014). International Migration of Couples. CESifo Working Paper No. 4927.

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

Kerr, S. P., Kerr, W. R., and Lincoln, W. F. (2015). Skilled Immigration and the Employment Structures of US Firms. *Journal of Labor Economics*, 33(S1):S147–S186.

Kerr, W. R. and Lincoln, W. F. (2010). The Supply Side of Innovation: H-1B Visa Reforms and U.S. Ethnic Invention. *Journal of Labor Economics*, 28(3):473–508.

Kleven, H. J., Landais, C., and Saez, E. (2013). Taxation and International Migration of Superstars: Evidence from the European Football Market. *American Economic Review*, 103(5):1892–1924.

Luxembourg Income Study (LIS) (2013). Database (multiple countries). http://www.lisdatacenter.org.

Moser, P., Voena, A., and Waldinger, F. (2014). German-Jewish Émigrés and US Invention. *American Economic Review*, 104(10):3222–3255.

OECD (2013b). OECD Education at a Glance 2013. OECD Indicators.

OECD (2013c). OECD Tax Statistics, Taxing Wages. http://dx.doi.org/10.1787/ctpa-twg-data-en.

OECD (2015). Benefits and Wages: Statistcs: Gross Replacement Rates (without Accounting for Taxes), Uneven Years 1961 to 2011. http://www.oecd.org/els/benefits-and-wages-statistics.htm.

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

Parey, M. and Waldinger, F. (2011). Studying Abroad and the Effect on International Labor Market Mobility. *Economic Journal*, 121(551):194–222.

Ramos, F. A. (1992). Outmigration and Return Migration of Puerto Ricans. In Borjas, G. J. and Freeman, R. B., editors, *Immigration and the Workforce: Economic Consequences for the United States and Source Areas*, chapter 2, pages 49–66. University of Chicago Press.

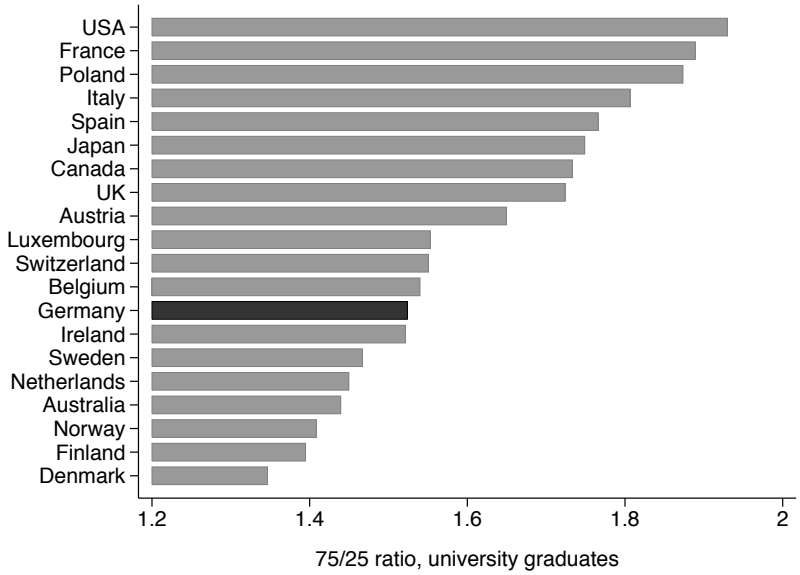
Roy, A. D. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers*, 3(2):135–146.

Ruggles, S., Alexander, J. T., Genadek, K., Goeken, R., Schroeder, M. B., and Sobeku, M. (2010). Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database].

Stolz, Y. and 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.

Widmaier, S. and Dumont, J.-C. (2011). Are Recent Immigrants Different? A New Profile of Immigrants in the OECD based on DIOC 2005/06. OECD Social, Employment and Migration Working Papers No. 126.

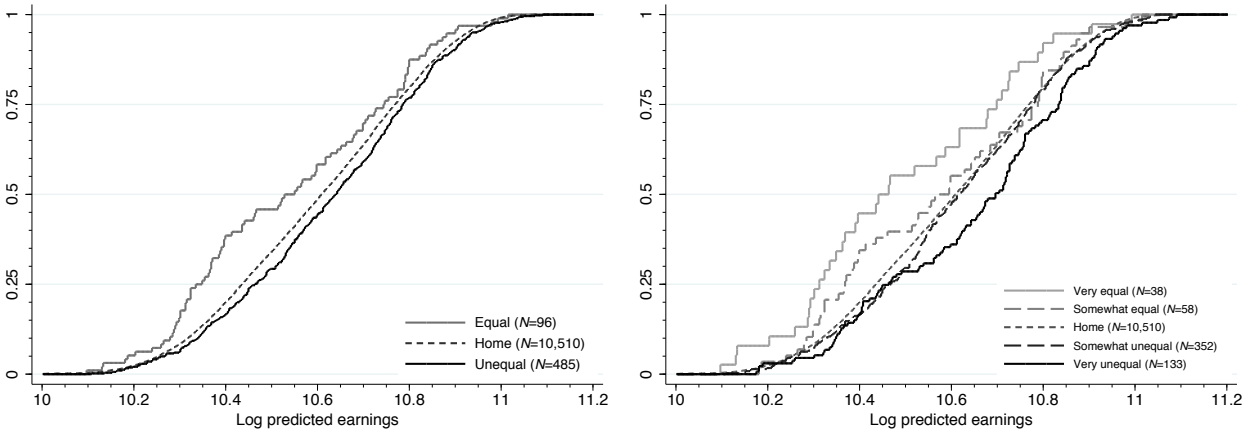
Figure 1: Earnings inequality among the high-skilled: Ratio of 75th to 25th percentile in the earnings distribution of university graduates



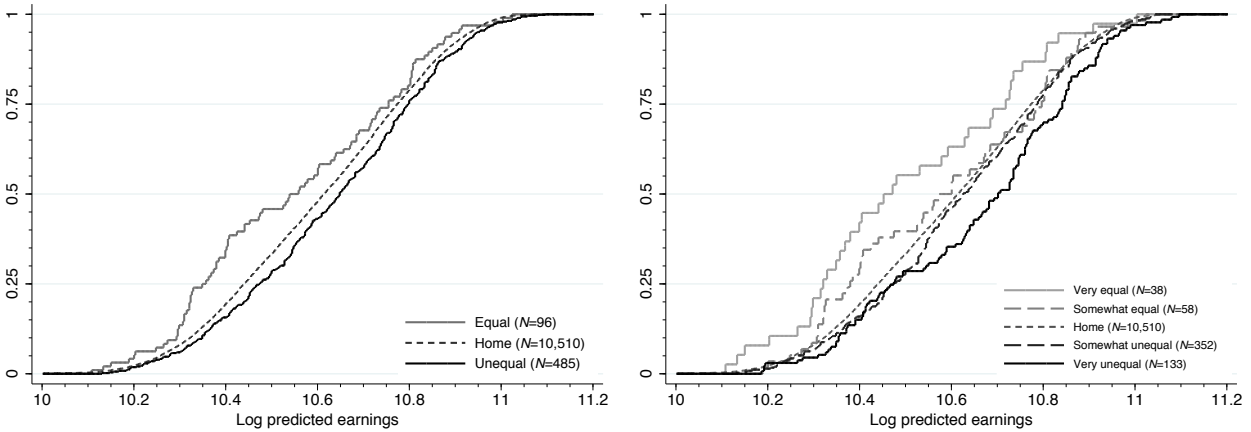
Notes: The figure shows the ratio of the 75th to the 25th percentile in the earnings distribution of university graduates. Authors' calculations based on country-specific earnings surveys (see Online Appendix Table A.3), showing averages over the period 1998 to 2010. Details on data sources and the construction of inequality measures are reported in section **??** and Data Appendix B.1.

Figure 2: Predicted earnings of migrants and non-migrants

(a) CDF for three and five groups of countries

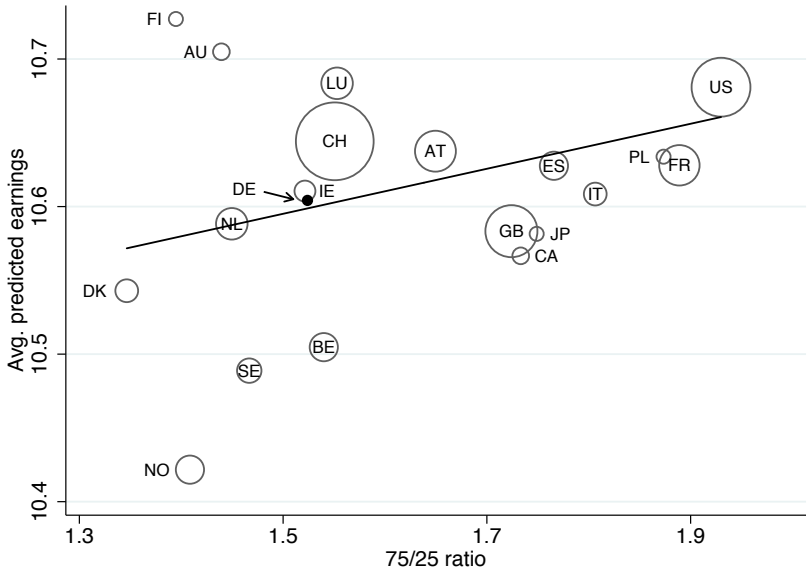


(b) CDF for three and five groups of countries – earnings prediction corrected for selection



Notes: Figure (a) shows CDFs of predicted earnings that are based on returns reported in column (1) of Table 2 for three groups: migrants to more equal countries, non-migrants, and migrants to less equal countries (left panel) and for five groups: migrants to very equal countries, migrants to somewhat equal countries, non-migrants, migrants to somewhat unequal countries, and migrants to very unequal countries (right panel). Figure (b) shows CDFs of predicted earnings that are based on selection-corrected returns reported in column (2) of Table 2 for the same groups of countries. Table 3 reports stochastic dominance tests. Online Appendix Figure A.2 shows kernel smoothed versions of the CDFs.

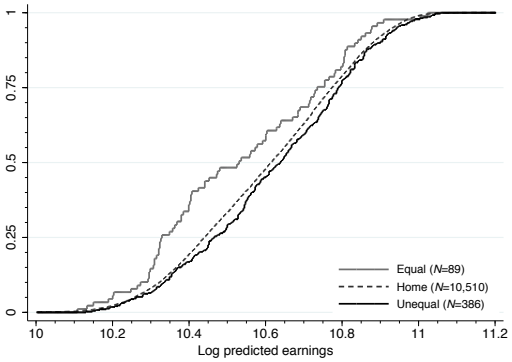
Figure 3: Predicted earnings and inequality across destinations



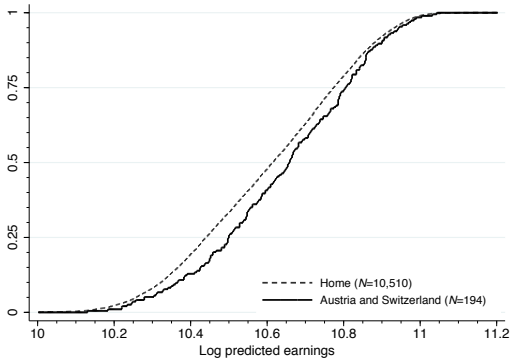
Notes: The figure shows average predicted earnings for migrants to each country and the corresponding 75/25 inequality ratio. Circle sizes are proportional to the number of migrants in each destination. The regression line reported in the figure is estimated in a weighted regression with weights equal to the number of migrants in each country. The slope coefficient is equal to 0.153 with a standard error of 0.081. An unweighted regression has a slope equal to 0.103 with a standard error of 0.101. For country labels see Data Appendix Table B.2.

Figure 4: Predicted earnings of migrants to Europe and Austria/Switzerland

(a) Europe



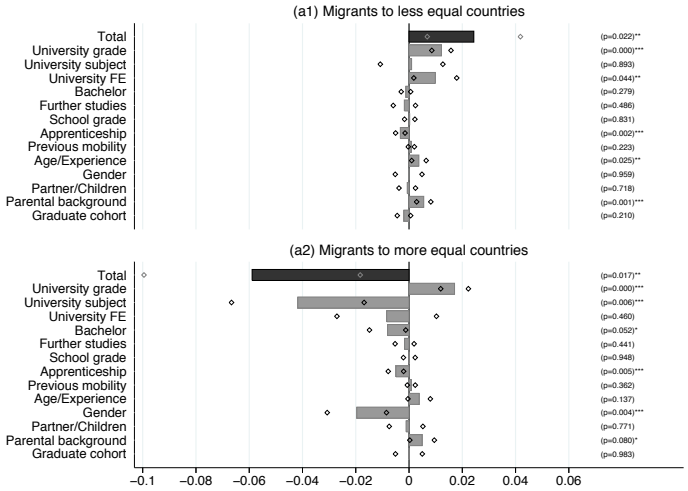
(b) Austria and Switzerland



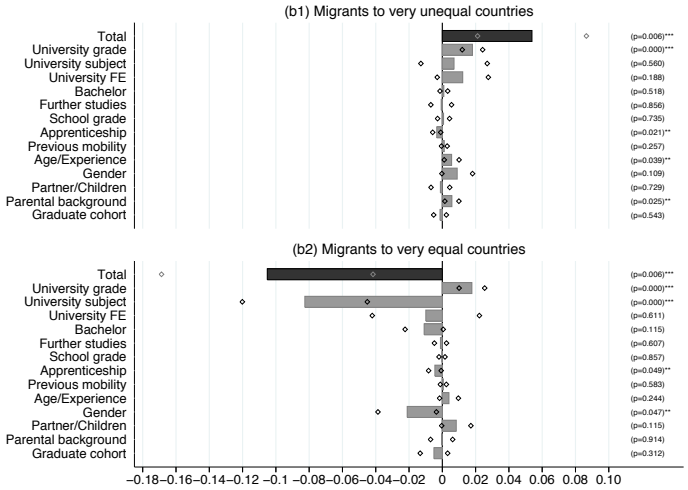
Notes: The figure shows CDFs of predicted earnings (prediction based on selection-corrected returns reported in column (2) of Table 2) for migrants to Europe (EU countries (2005), Norway, and Switzerland) and non-migrants in panel (a); and to Austria or Switzerland and non-migrants in panel (b). Online Appendix Table A.9 (panels A and B) reports stochastic dominance tests.

Figure 5: Decomposition of predicted earnings

(a) Migrants to less equal and more equal countries



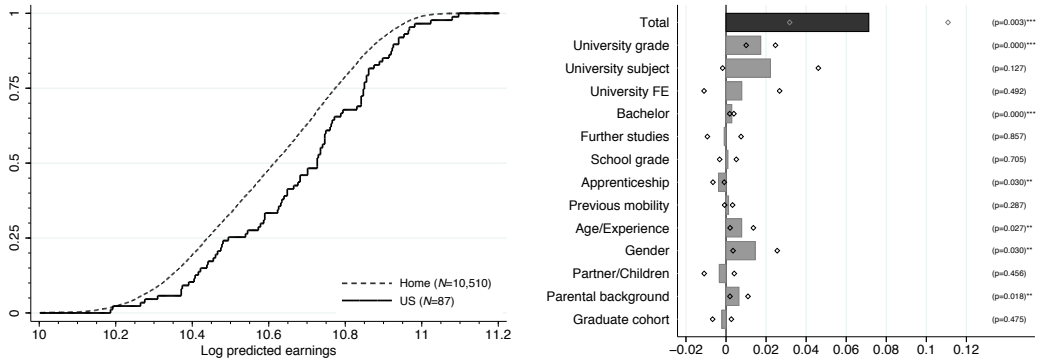
(b) Migrants to very unequal and very equal countries



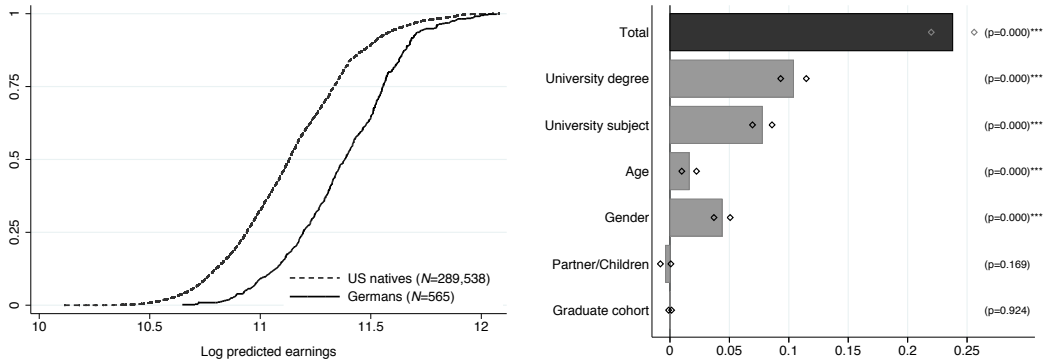
Notes: Panel (a1) decomposes the mean difference in predicted earnings between migrants to less equal countries and non-migrants. The top bar (black) measures the total difference in predicted earnings. The other bars decompose the total difference into the contributions of groups of characteristics (e.g. university grade). More specifically, the size of the gray bars in panel (a1) is obtained by multiplying estimated returns ($\hat{\beta}_k^{Home}$) for non-migrants from column (1) in Table 2 (where k indexes a group of characteristics, e.g. all parental background variables or all university fixed effects) with average characteristics of migrants to less equal countries ($\overline{x}_k^{Less\ equal}$) and average characteristics of non-migrants ($\overline{x}_k^{Home}$), and then subtracting $\hat{\beta}_k^{Home}\overline{x}_k^{Home}$ from $\hat{\beta}_k^{Home}\overline{x}_k^{Less\ equal}$. Panel (a2) presents the equivalent decomposition between migrants to more equal destinations and non-migrants. Panel (b) presents corresponding results to very unequal and very equal countries. Diamonds indicate 90% confidence intervals. Confidence intervals and p-values are obtained from bootstrapped standard errors (based on 4,999 replications). Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Figure 6: Predicted earnings of migrants to the United States

(a) based on returns in German sample (DZHW data)



(b) based on returns in U.S. sample (ACS data)



Notes: Figure (a): The left panel shows CDFs of predicted earnings (prediction based on selection-corrected returns reported in column (2) of Table 2) for migrants to the United States and for non-migrants. Online Appendix Table A.9 (panel C) reports stochastic dominance tests. The right panel decomposes the mean difference in predicted earnings between migrants to the United States and non-migrants. The top bar (black) measures the total difference in predicted earnings. The other bars decompose the total difference into the contributions of groups of characteristics (e.g. university grade). Figure (b): The left panel shows CDFs of predicted earnings in the United States. Prediction based on coefficients of the Mincer regression reported in Online Appendix Table A.10 (column (1)) using American Community Survey (ACS) data on U.S. natives. The right panel shows a decomposition of predicted earnings that decomposes the mean difference in predicted earnings between German migrants to the United States and U.S. natives. Diamonds indicate 90% confidence intervals. Confidence intervals and p-values are obtained from bootstrapped standard errors (based on 4,999 replications). Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table 1: Summary statistics for German university graduates

	Full sample		Working in Germany	Abroad more equal	Abroad less equal
	Mean	SD	Mean	Mean	Mean
Job characteristics (after five years)					
Working abroad	0.052	–	0	1	1
Annual earnings in Euro (2001 prices)	43,491	19,334	43,265	39,458	49,231
Potential experience in months	69.201	4.161	69.197	69.719	69.194
Postgraduate education					
PhD completed	0.191	–	0.182	0.313	0.371
Further (non-PhD) degree completed	0.073	–	0.071	0.125	0.122
Education first degree					
Final university grade	2.018	0.681	2.032	1.698	1.787
Studying abroad	0.078	–	0.072	0.240	0.169
Age at graduation	26.994	2.664	27.026	26.271	26.437
ERASMUS/Total students in subject	0.040	0.057	0.039	0.052	0.050
Education before first degree					
Final school grade	2.110	0.639	2.119	1.951	1.959
Apprenticeship	0.220	–	0.225	0.094	0.138
Previous mobility					
Studied in same state as high school	0.659	–	0.663	0.583	0.581
Personal characteristics					
Female	0.445	–	0.444	0.594	0.445
Partner	0.780	–	0.782	0.740	0.736
Married	0.416	–	0.421	0.281	0.344
Child(ren)	0.291	–	0.297	0.156	0.184
Parental background					
Mother's education (years)	13.459	3.102	13.423	14.458	14.035
Father's education (years)	14.852	3.065	14.816	15.458	15.493
Mother self-employed	0.092	–	0.093	0.063	0.091
Mother salaried employee	0.597	–	0.596	0.677	0.619
Mother civil servant	0.108	–	0.105	0.177	0.148
Mother worker	0.100	–	0.103	0.042	0.049
Mother did not work	0.103	–	0.104	0.041	0.093
Father self-employed	0.194	–	0.191	0.188	0.262
Father salaried employee	0.447	–	0.448	0.479	0.406
Father civil servant	0.223	–	0.221	0.271	0.258
Father worker	0.113	–	0.116	0.063	0.062
Father did not work	0.023	–	0.024	0.000	0.012
Observations	11,091		10,510	96	485

Notes: The table shows summary statistics of German university graduates at five years after graduation. Information on earnings is available for 10,315 of the 11,091 graduates. Average annual earnings of 43,491 Euros in 2001 prices correspond to around 79,084 U.S. dollars in 2014 prices.

Table 2: Augmented Mincer regression for university graduates in Germany

	(1)		(2)		(3)	
Dependent variable	Labor earnings		Labor earnings		Working in Germany	
	OLS		Heckman sel. model		Selection equation	
Education first degree						
Final university grade	0.048*	(0.027)	0.046*	(0.027)	0.079	(0.203)
Final grade squared	-0.023***	(0.006)	-0.023***	(0.006)	-0.007	(0.048)
Bachelor's degree	-0.131***	(0.028)	-0.132***	(0.028)	0.049	(0.158)
Age at graduation	-0.026**	(0.011)	-0.026**	(0.011)	-0.013	(0.097)
Age squared	0.000*	(0.000)	0.000*	(0.000)	0.001	(0.002)
Postgraduate education						
PhD completed	-0.003	(0.011)	0.000	(0.013)	-0.367***	(0.065)
Further (non-PhD) degree completed	-0.024	(0.015)	-0.021	(0.016)	-0.251***	(0.085)
Education before first degree						
Final school grade	-0.041	(0.034)	-0.043	(0.034)	0.109	(0.224)
School grade squared	0.009	(0.008)	0.010	(0.008)	-0.011	(0.052)
Apprenticeship	0.037***	(0.010)	0.037***	(0.010)	0.078	(0.071)
Previous mobility						
Studied in same state as high school	-0.010	(0.008)	-0.012	(0.008)	0.131***	(0.049)
Potential work experience						
Experience in months	-0.058***	(0.022)	-0.059***	(0.022)	0.096	(0.138)
Experience squared	0.000***	(0.000)	0.000***	(0.000)	-0.001	(0.001)
Personal characteristics						
Female	-0.131***	(0.008)	-0.131***	(0.008)	-0.047	(0.053)
Partner	0.066***	(0.009)	0.065***	(0.009)	0.070	(0.058)
Married (additionally)	0.028***	(0.009)	0.028***	(0.009)	0.027	(0.058)
Child(ren)	-0.040***	(0.009)	-0.041***	(0.010)	0.210***	(0.065)
Parental background						
Mother's education (years)	0.003*	(0.002)	0.003*	(0.002)	-0.002	(0.010)
Father's education (years)	0.003*	(0.002)	0.003*	(0.002)	-0.019*	(0.010)
Mother self-employed	-0.008	(0.017)	-0.009	(0.017)	0.107	(0.112)
Mother salaried employee	-0.012	(0.013)	-0.012	(0.013)	0.010	(0.086)
Mother civil servant	-0.019	(0.018)	-0.019	(0.017)	-0.013	(0.112)
Mother worker	-0.001	(0.016)	-0.003	(0.016)	0.194	(0.122)
Father self-employed	0.054**	(0.025)	0.056**	(0.025)	-0.260	(0.195)
Father salaried employee	0.041*	(0.024)	0.041*	(0.024)	-0.053	(0.192)
Father civil servant	0.027	(0.025)	0.028	(0.025)	-0.132	(0.196)
Father worker	0.003	(0.026)	0.003	(0.026)	0.009	(0.209)
ERASMUS places/students					-1.197***	(0.424)
Mills ratio			-0.050	(0.095)		
Graduate cohort FE	YES		YES		YES	
Subject FE	YES		YES		YES	
University FE	YES		YES		YES	
R-sq./Pseudo R-sq.	0.282				0.132	
Observations	9,778		9,778		10,315	

Notes: Column (1) reports results from the augmented Mincer regression. Column (2) reports results from the augmented Mincer regression that controls for selection in the decision to work in Germany using a Heckman selection correction. Column (3) reports the corresponding selection equation, which predicts working in Germany with the number of ERASMUS places normalized by the cohort size in a graduate's university department. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Stochastic dominance tests

	OLS			Heckman selection correction		
		p-value			p-value	
	Test statistic	Kolmogorov-Smirnov	Barrett-Donald	Test statistic	Kolmogorov-Smirnov	Barrett-Donald
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Selection to *more equal* and to *less equal* destinations						
'Equal' vs 'Home'	0.187	0.001 ***	0.006 ***	0.182	0.002 ***	0.022 **
'Home' vs 'Unequal'	0.061	0.031 **	0.098 *	0.071	0.009 ***	0.083 *
'Equal' vs 'Unequal'	0.220	0.000 ***	0.001 ***	0.218	0.000 ***	0.004 ***
Panel B: Selection to *very equal*, to *somewhat equal*, to *somewhat unequal*, and to *very unequal* destinations						
Stochastic dominance tests for very equal and very unequal destinations						
'Very equal' vs 'Home'	0.258	0.007 ***	0.018 **	0.249	0.009 ***	0.041 **
'Home' vs 'Very unequal'	0.144	0.004 ***	0.017 **	0.162	0.001 ***	0.014 **
'Very equal' vs 'Very unequal'	0.301	0.005 ***	0.008 ***	0.301	0.005 ***	0.012 **
Stochastic dominance tests for more similar destinations						
'Very equal' vs 'Somewhat equal'	0.179	0.231	0.379	0.196	0.171	0.310
'Somewhat equal' vs 'Home'	0.147	0.083 *	0.177	0.142	0.099 *	0.171
'Home' vs 'Somewhat unequal'	0.057	0.109	0.235	0.065	0.055 *	0.173
'Somewhat unequal' vs 'Very unequal'	0.133	0.033 **	0.101	0.136	0.027 **	0.096 *

Notes: The table reports one-sided Kolmogorov-Smirnov test statistics and Kolmogorov-Smirnov and Barrett and Donald p-values for CDFs in Figure 2. Barrett and Donald p-values are bootstrapped, following equation (11) in Barrett and Donald (2003, p. 82). In the top row ('Equal' versus 'Home'), we test the null hypothesis that the CDF of migrants to more equal destinations stochastically dominates the CDF of non-migrants, and similarly for other rows. The bootstrap is based on 4,999 replications. See text for details. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Cross-country regressions

	(1)	(2)	(3)	(4)	(5)	(6)
75/25 ratio	0.153*	0.180***	0.174**	0.110*	0.247***	0.147*
	(0.081)	(0.058)	(0.077)	(0.057)	(0.081)	(0.071)
Mean earnings		0.110***				0.102*
		(0.033)				(0.056)
Tertiary-educated unemployment share			-0.007			0.005
			(0.007)			(0.009)
Family expenditure				-0.023*		-0.012
				(0.011)		(0.011)
Life satisfaction					0.050*	0.003
					(0.024)	(0.036)
Constant	10.366***	9.161***	10.353***	10.484***	9.849***	9.281***
	(0.144)	(0.413)	(0.138)	(0.104)	(0.276)	(0.531)
R-sq.	0.183	0.475	0.204	0.317	0.282	0.514
Observations	19	19	19	19	19	19

Notes: The table reports weighted regressions of average predicted earnings of migrants in each country on the corresponding 75/25 ratio and potential confounders. See Data Appendix B.2 for details on data sources and Data Appendix Table B.2 for country data. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Summary of decomposition results

	(1)	(2)	(3)	(4)	(5)
	Less equal destinations	More equal destinations	Very unequal destinations	Very equal destinations	United States
Total	consistent	consistent	consistent	consistent	consistent
University grade	consistent	reject	consistent	reject	consistent
University subject	–	consistent	–	consistent	–
University fixed effect	consistent	–	–	–	–
Bachelor	–	–	–	–	consistent
Further studies	–	–	–	–	–
School grade	–	–	–	–	–
Apprenticeship	reject	consistent	reject	consistent	reject
Previous mobility	–	–	–	–	–
Age/Experience	consistent	–	consistent	–	consistent
Gender	–	consistent	–	consistent	consistent
Partner/Children	–	–	–	–	–
Parental background	consistent	–	consistent	–	consistent
Graduate cohort	–	–	–	–	–

Notes: The table summarizes results from the Blinder-Oaxaca decomposition shown in Figure 5 and 6(a). 'Consistent' indicates that the selection along the corresponding characteristic is significantly different from 0 at a 5% level of significance and in line with the model prediction. 'Reject' indicates that the selection along the corresponding characteristic is significantly different from 0 at a 5% level of significance, in the direction not in line with the model prediction.