I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10891

Limits to Wage Growth: Understanding the Wage Divergence between Immigrants and Natives

Apoorva Jain
Klara Sabirianova Peter

JULY 2017

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10891

Limits to Wage Growth: Understanding the Wage Divergence between Immigrants and Natives

Apoorva Jain
University of North Carolina-Chapel Hill

Klara Sabirianova Peter
University of North Carolina-Chapel Hill, IZA and CEPR

JULY 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Limits to Wage Growth: Understanding the Wage Divergence between Immigrants and Natives*

This study finds evidence of wage divergence between immigrants and natives in Germany using a country-wide household panel from 1984 to 2014. We incorporate the possibility of wage divergence into a two-period model of economic assimilation by modeling the differences in the efficiency of human capital production and prices per unit of human capital between immigrants and natives. Individual rates of wage convergence are found to be higher for immigrants who fled warfare zones, belong to established ethnic networks, and acquired more years of pre-migration schooling. Using a doubly robust treatment effect estimator and the IV method, the study finds that the endogenous post-migration education in the host country contributes substantially to closing the wage gap with natives. The treatment effect is heterogeneous, favoring immigrants who are similar to natives. This paper also addresses the commonly ignored sample selection issue due to non-random survey attrition and employment participation. Empirical evidence favors the "efficiency" over the "discrimination" channels of wage divergence.

JEL Classification: J15, J24, J31, J61, F22, I26

Keywords: migration, assimilation, divergence, wage growth, skill prices, post-migration human capital, discrimination, doubly robust estimator, instrumental variables, panel, Germany

Corresponding author:
Klara Sabirianova Peter
Department of Economics
Carolina Population Center
University of North Carolina-Chapel Hill
Chapel Hill, NC 27599
USA

E-mail: kpeter@unc.edu

* We are thankful to Charles Becker, Luca Flabbi, Donna Gilleskie, Jane Cooley Fruehwirth, Krista Perreira, and Helen Tauchen for useful comments. We also acknowledge helpful comments received from participants of the UNC Applied Microeconomics Workshop and the 12th International German SOEP User Conference (Berlin, 2016). We are grateful to the Carolina Population Center and its NIH Center grant (P2C HD050924) for general support.

1. Introduction

The global economic integration and the string of recent humanitarian crises around the world have spurred a great deal of immigrant movement. While many governments have been accepting immigrants, there has been some opposition based on concerns around how well the immigrants are going to assimilate into their new society. Existing theoretical models of immigrant human capital investment generally predict that immigrants will experience faster wage growth than comparable natives due to their lower cost of investment in human capital and greater incentives to acquire more skills (Chiswick, 1978; Duleep and Regets, 1999; Borjas, 1999). This prediction has been put to test by numerous studies, which typically found a fairly rapid rate of wage convergence. After accounting for non-random out-migration and immigrant cohort quality, the degree of wage convergence becomes not as fast as it was previously believed, but it remains positive (Borjas, 1995, 2015; Lubotsky, 2007).

This paper shows that positive post-migration wage growth and relatively large returns to an additional year of stay in the host country do not necessarily result in wage convergence between immigrants and natives. Using the German Socio-Economic Panel (GSOEP) from 1984 to 2014, we find substantial evidence of an increasing native-immigrant wage gap over an individual's life-cycle.[1] Not only is the average wage of immigrants smaller than that of natives, the rate of wage growth is also significantly lower for immigrants compared to the native population, which is counter to the existing economic models. Currently, there is no methodological framework that helps to explain the wage divergence between immigrants and comparable natives. This study attempts to fill this gap.

We extend the standard theoretical model of immigrant economic assimilation to allow for the possibility of wage divergence and derive testable hypotheses. Our theoretical model shows that wage divergence is possible if natives are relatively more efficient in the production of human capital and/or if the price per unit of human capital increases over the life-cycle at a higher rate for natives than for immigrants. We preserve the key features of Borjas's (1999) framework by allowing post-migration human capital accumulation to vary with the level of pre-existing human capital, skill transferability, and the discounting factor of future earnings.

[1] Evidence of wage divergence between immigrants and natives is also found in Italian data by Venturini and Villosio (2008) and in the GSOEP by Zibrowius (2012). Both studies infer wage divergence from lower wage returns to work experience for immigrants compared to natives. We define wage divergence differently as a slower wage growth of immigrants relative to the wage growth of comparably skilled natives who are at the same point in the life cycle.

We also develop and estimate an empirical model of wage convergence. The dependent variable in this model is the average annual growth in relative wage over the five-year period. Relative wage shows the position of an immigrant in the wage distribution of comparable natives of the same age, schooling, and location type. This model was inspired by the wage growth equation estimated in Borjas (2015). The dependent variable in his paper is assimilation rates aggregated from U.S. Census data and measured as the 10-year wage growth experienced by an immigrant cohort from a given country of origin relative to the wage growth experienced by comparably aged native workers. Unlike the cohort-level approach in Borjas's study, our rates of assimilation are individual-specific and allowed to vary with individual characteristics at arrival, post-migration investment, and characteristics of the home country at the time of entry. One of the advantages of the individual-level wage growth model is that it accounts for permanent unobserved individual heterogeneity, yet also allows for estimating the effect of time-constant factors on wage growth.

In addition to using a long-term panel of individuals, this study benefits greatly from the life history calendars provided by the GSOEP for each immigrant between the ages of 15 and 65. Based on the calendar data, we construct more accurate measures of pre- and post- migration schooling and job training. Such information is rarely available in other datasets. Most of the previous papers use highly crude measures of pre- and post-education based on total years of schooling and age-at-migration. As a result, these measures suffer from measurement error, and using them can also generate systematic bias (Duleep, 2015).

One important concern with including post-migration accumulation of human capital in the wage convergence model is the potential endogeneity of new skill acquisition. Due to the inherent difficulty of dealing with endogeneity, this issue has been largely avoided in the migration literature. Skuterud and Su (2015) is the only study we are aware of that attempts to address the endogeneity of post-migration schooling by including individual fixed effects in the wage-level equation. Even though permanent individual heterogeneity is accounted for in the growth equation, the theoretical model predicts that individual decisions about new skill acquisition may be based on anticipated wage gains, and thus new investment could be endogenous in the growth equation. In identifying the treatment effect of post-migration education, we rely on the lagged investment variable, the selection-on-observables under the conditional mean independence assumption, and instrumental variables (IVs). For IVs, we employ demand-supply shifters in government-sponsored training programs and potential schooling interruptions due to wars and internal conflicts in the country of origin during early schooling age from 6 to 10. Using a doubly robust treatment effect estimator and the IV-LATE method, the study finds that the endogenous post-migration education in the host country contributes substantially to closing the wage gap with natives, and that this contribution greatly exceeds the positive convergence effect of pre-migration investment.

We recognize that potential selection bias could be a problem in estimating the wage growth model despite the growth-specification's advantage over the level-specification in accounting for permanent individual heterogeneity in selection. Generally, growth variables cause a greater loss of valid observations in the estimation sample. For example, we use the minimum of three non-missing data points in calculating the average relative wage growth over a 5-year period. Missing data on growth rates may be related to out-migration, respondent's death, non-response at follow-up, exit from employment between the two survey rounds, or wage non-reporting conditional on being employed.[2] Using the Heckman-style correction and inverse propensity weighting procedures, our analysis shows that unobserved growth rates from all the above sources of missing data do not contribute to the divergence of relative wages.

Our estimates reveal that the rates of wage divergence tend to be higher among immigrants who are males, have less educated parents, are not ethnic Germans, acquire fewer years of formal schooling in the home country, come from lower-income countries, and are part of smaller ethnic networks. Immigrants who escaped political violence in their home country have higher assimilation rates on average. Compared to the U.S., where the average rate of economic assimilation is declining with time (Borjas, 2015), Germany has a marginally upward trend in wage convergence over calendar time.

Only one of the two wage divergence channels conjectured by the theoretical model – namely, higher efficiency of natives in the production of human capital– is consistent with the data. Considerably higher divergence rates during the investment period compared to the later stages of working career are in line with the efficiency story. We find that immigrants with fewer linguistic and cultural barriers benefit the most from host-country education in terms of the future wage trajectory. At the same time, immigrants who are distant from natives in observed characteristics have a small and statistically insignificant effect of post-migration education on wage convergence. These

[2] In the "level" specification of the wage assimilation model, several studies addressed selection bias due to non-random out-migration and panel attrition (e.g., Bellemare, 2007; Constant and Massey, 2003; Dustmann and Glitz, 2011; Dustmann and Görlach, 2016; Fertig and Schurer, 2007). Employment-related sources of selection bias have been generally overlooked in the migration literature, although separate estimates of employment assimilation rates are common (see review of studies in Kerr and Kerr, 2011).

two findings also support the efficiency explanation for observed wage divergence. The second channel of wage divergence – differential change in skill prices favoring natives – is not supported by the available data. In fact, the perceived discrimination against immigrants weakens with age, while the native-immigrant wage gap moves in the opposite direction.

The rest of the paper proceeds as follows. Section 2 presents a set of empirical facts concerning the economic assimilation of immigrants in Germany. Section 3 develops a simple theoretical model of wage growth that highlights the main channels behind economic convergence/divergence in wages between immigrants and natives. Section 4 discusses the empirical strategy for estimating the model of wage convergence, with a special emphasis on both measuring the factors of economic assimilation and addressing the selectivity and endogeneity issues. Section 5 presents model estimates, including reduced-form equations for wage convergence and wage growth, estimated treatment effects of post-migration human capital, as well as empirical evidence for channels of divergence. Finally, Section 6 concludes.

2. Empirical Evidence on the Economic Assimilation of Immigrants

In this section, we present a set of empirical facts concerning the economic assimilation of immigrants in Germany. The facts are drawn from the statistical comparison of labor market outcomes between immigrants and natives using the GSOEP, an annual panel of households from 1984 to 2014.[3] Since the GSOEP survey is well documented and widely used, we provide a data description in Appendix A1 instead of the main text. The immigrant status is defined based on the country of birth outside either East or West Germany. For all analyses, we limit the sample to those who were between the ages of 17 and 65 at the time of survey and who reside in West Germany.[4] The sample of immigrants is further constrained to those who arrived in Germany after 1960 at age 15 or older. Child immigrants are excluded from the analysis because pre-migration

[3] GSOEP is a very popular data source in the migration literature, as it is one of a few national longitudinal surveys with a large representation of immigrants. Dustmann and Görlach (2015) highlight several advantages of longitudinal datasets over frequently used synthetic cohorts, repeated cross-sections, and retrospective panels on earnings linked to a single cross-section of households. The main advantage is the unbiased identification of immigrant assimilation profiles conditional on proper modeling of the non-random selection into employment and out-migration.
[4] Less than two percent of all immigrants in population reside in East Germany. It is common in the migration literature based on GSOEP to exclude this subsample from the analysis (e.g., Basilio *et al*., 2017).

history records begin from age 15. Our estimation sample includes 31,215 natives and 7,496 immigrants.

2.1 Sample composition

From the sample composition of immigrants and natives shown in Table 1, we see that the immigrant sample on average has a higher share of females by about 2 percentage points, is 4 years older, acquired one year less schooling, and has considerably less educated parents than the native population sample does. 86 percent of surveyed immigrants but only 75 percent of natives reside in urban areas. All the above mentioned mean differences between the two samples are statistically significant at the 1 percent level. An average adult immigrant arrives to Germany at age 26 with 10 years of formal schooling and about 1 year of previous job training and spends 19 years in the host country. After migration, only 14 percent receive formal schooling in Germany and 14 percent acquire job training, with some overlap. In total, about 22 percent of immigrants study in Germany.

The composition of immigrants in the GSOEP reflects German migration history. Before the unification of Germany in 1990, top-sending countries were countries that signed guest-worker recruitment agreements with Germany in the 1960s: Turkey (26 percent of the sample of immigrants), Yugoslavia (15 percent), and Italy (13 percent). Immigrants from Poland also had a large share (9 percent) due to the influx of Polish refugees in the 1980s. After 1990, German migration policy has shifted from guest-worker programs and family reunification to programs of resettlement of ethnic Germans mainly from the former Soviet Union and East Europe. As a result, ethnic Germans from Russia and Kazakhstan took the top two spots among new arrivals (16 and 14 percent, respectively). Shortly after German reunification, the number of refugees climbed sharply, triggered by the Yugoslav wars, perpetual series of wars in the Middle East, and other international conflicts. As a result, the share of immigrants from ex-Yugoslavia remains large even after 1990 (11 percent). The overall share of immigrants from the Middle East and North Africa (excluding Turkey) is about 5 percent of the post-1990 arrival cohort, but this share is expected to rise in light of the current migration crisis in Europe.[5]

[5] In calculating the percent shares in this paragraph, the longitudinal data is collapsed such that each immigrant is counted once. The compositions of immigrants by year and county of birth in the GSOEP sample and official population statistics are highly correlated (0.83). Some mismatch that arises due to the idiosyncratic sampling of immigrants in the GSOEP is adjusted by using probability sampling weights (see Appendix A1 for further details).

2.2 Labor market outcomes of immigrants and natives

Table 2 reports unconditional and "conditional on common covariates" mean differences in labor market outcomes between immigrants and natives. Each column represents one of three labor market outcomes: the real hourly wage, the probability of being employed, and the probability of being unemployed conditional on being in the labor force.[6] First, note that due to late arrivals, the average immigrant enters the estimation sample at an older age than the average native does. As a result, the comparison of unconditional outcomes between the two groups could be misleading. The raw sample means show that immigrants earn a 6.5-percent higher hourly wage than natives do. However, once the age is fixed using a flexible quartic polynomial function, the wage gap between immigrants and natives turns out to be substantial. On average, immigrants earn an hourly wage that is 17 percent less than a comparably aged native worker. The gap narrows to 11 percent once we control for other observed characteristics such as gender, years of schooling, urban current residence, and year fixed effects, but it remains sizeable. Employment outcomes, even unconditional ones, are also considerably worse for immigrants than natives. In the raw data, immigrants have an 8.5 percentage-point lower employment participation rate and an 8 percentage-point higher unemployment rate than natives do. The conditional native-immigrant gap is about 10 percentage points in employment participation and 7 percentage points in the unemployment rate.

2.3 Wage returns to the length of stay since migration

When measuring the economic assimilation of immigrants in terms of their wage trajectory in the host country, it is important to distinguish between the post-migration wage progression relative to the immigrant's own entry wage and the wage progression of immigrants relative to natives; see Borjas (1999) for the discussion of two alternative definitions of economic assimilation. These two definitions are associated with two different concepts of wage convergence. The first concept, which is analogous to the beta-convergence in the macro growth literature, implies that the wages of immigrants with low and high unobserved skills move towards each other when immigrants with a lower entry wage (as a proxy for unobserved skills) have faster post-migration wage growth. The second concept of wage convergence implies that the wages of immigrants are catching up with the wages of comparably skilled natives as immigrants spend more time in the host country. This second concept is the focus of our study.

[6] Further details on how each outcome is constructed are provided in Appendix A2.

Within the first conceptual framework, the average assimilation rate is typically obtained as the slope coefficient on the number of years since migration in a standard wage equation estimated over a sample of immigrants. By allowing the unobserved individual heterogeneity to influence both the random intercept and random slope, we can test for the presence of conditional wage convergence between low- and high-skill immigrants, as shown below.

$$w_{it} = a_0 + (\bar{\delta} + b_i)YSM_{it} + \gamma_X X_{it} + \gamma_\varphi \varphi(AGE_{it}, TIME_{it}) + a_i + \varepsilon_{it}, \qquad (1)$$

where w_{it} is the log of hourly wage of individual i at survey time t; YSM_{it} is the number of years since migration; X_{it} is the vector of observed individual characteristics; $\varphi(AGE_{it}, TIME_{it})$ denotes a flexible function of the immigrant age and survey time; $\bar{\delta}$ is the average wage return on spending an additional year in the host country; b_i is the individual-specific deviation from the average rate of assimilation with zero mean; a_i is a random intercept capturing immigrants' unobserved skills with zero mean; $\varepsilon_{is} \sim N(0, \sigma_\varepsilon^2)$ is an i.i.d. error independent of a_i's and b_i's. Equation (1) belongs to the class of linear mixed-effects models with correlated random intercepts and slopes. In the mixed model, a_i and b_i are assumed to be drawn from a joint bivariate normal distribution with mean zero and a variance-covariance matrix with elements $\sigma_a, \sigma_b,$ and σ_{ab}. A negative covariance between the two random effects ($\sigma_{ab} < 0$) implies that immigrants with lower unobserved skills have a faster rate of wage assimilation, holding observed characteristics constant. Hence, a negative correlation sign, if found, would support the hypothesis of conditional convergence between low- and high-skill immigrants of similar observed characteristics.

We draw the distribution of the estimated returns to a year of stay in Germany in Figure 1. The mean return is about 1 percent in annual wage gains. This estimate is close to previously reported estimates for Germany (Basilio *et al.*, 2009; Basilio *et al.*, 2017). Beyond the mean estimates, we find significant heterogeneity in individual rates of the immigrant's wage progression; see the left panel of Figure 1. Six percent of all immigrants experience an average decline in their real wage over the life cycle. Figure 1 also depicts a strongly negative correlation between the best linear unbiased predictors of $\hat{a}_i$ and $\hat{b}_i$ (-0.76). These estimates are obtained from a simplified mixed-effects wage model with an abbreviated list of controls.[7] Jain and Peter (2017) use a more refined joint hazard-longitudinal (JHL) model that accounts for the endogenous timing of migration, non-

[7] Full estimates of Equation (1) including its OLS specification are reported in Appendix Table W1. We use the same covariates as in Table 2 plus fixed effects for the country of origin.

random attrition, and the selection into employment. The JHL model finds the average assimilation rate in Germany for the same period to be lower, at about 0.7 percent increase in wage per each additional year of stay. Yet, this return is substantial considering that the average adult immigrant spends almost 20 years in the host country. The JHL model also finds the inverse relationship between unobserved skills and the rate of wage assimilation (the coefficient of correlation is -0.83). However, as we show below, the type-I wage convergence between low- and high-skill immigrants does not imply that wages of immigrants as a group are converging to the wage level of their native counterparts.

2.4 Age profiles of relative wages: first evidence of wage divergence

A positive wage return on years since migration is a necessary but not sufficient condition for the successful economic assimilation of immigrants. Wage convergence between immigrants and natives is not going to be achieved if the wages of natives grow at a faster rate than the wages of immigrants. Let's compare the life-cycle trajectories of the log hourly wage between immigrants and natives, shown in Panel A of Figure 2. Consistent with the positive return to the length of stay in the host country, immigrants' wages increase over the life-cycle. Yet, natives have a much steeper age-wage profile and thus higher rates of wage growth compared to the immigrant population (at least until about age 50). Wage trajectories are striking and somewhat unexpected in that they are indicative of diverging wage trajectories between immigrants and natives. The "catching-up" effect found in some U.S. studies (Borjas, 1999) does not show up in this figure.

In Figure 2A, we also observe that the average wage is higher for immigrants than natives in the early work career phase. This result could simply reflect the compositional differences between immigrants and natives. Indeed, once we control for basic observed characteristics, the wage gap favoring younger immigrants vanishes, as we see from the life-cycle trajectory of relative wage in Figure 2D.

In constructing relative wages, we first obtain the percentile values of the residuals from the regression of native wages on the X vector in year t. Then, we predict residuals for each immigrant and find the corresponding percentile θ_{it} in the residual distribution of natives. Using this method, we obtain three measures of relative wages depending on the specification of the X vector: (i) unconditional if X includes only the intercept; (ii) age-specific if X also contains a quartic polynomial in age; and (iii) conditional if X includes the level of schooling and urban residence in addition to the intercept and a quartic polynomial in age. In the latter case, θ_{it} is interpreted as the position of the immigrant in the wage distribution of comparable natives with the same observed characteristics.

In Figure 2, we plot all three measures of relative wage over the life-cycle. The unconditional relative wage closely follows the trajectory of the hourly wage for both immigrants and natives. By construction, the native trajectory of age-specific and conditional wage lies around the 50^{th} percentile line. The small deviation arises from the parametric function of age and aggregation. If we only condition on age, as shown in Panel C, the wage gap between immigrants and natives is about 4 percentiles at age 25, but it rapidly increases and reaches a 14 percentile-difference by age 50. If we also control for schooling and location, as in Panel D, the wage gap is noticeably smaller; it is even close to nil during the early work career, but it widens to a substantial 8-9-percentile difference for ages 45 to 60. In other words, despite a solid increase in wages after migration, the position of immigrants in the wage distribution of comparable natives falls with age.[8]

2.5 Selection into employment and survey participation

In Figure 3, we illustrate the life-cycle trajectories in employment outcomes. Similar to wage differentials, there is a considerable native-immigrant gap in employment participation rates (about 13 percentage points at age 40) and unemployment rates conditional on being in the labor force (6 percentage points at age 40). The gap is also large for the probability of exiting employment conditional on working in previous year; 6.6 percent of immigrant workers and only 3.4 percent of native-born workers at age 40 lose their job annually. Trajectories in unemployment probabilities show no sign of convergence. However, the gap in employment participation and exit rates seems to be closing over time and achieving convergence by the end of working career. The convergence in employment probabilities and divergence in unemployment probabilities may co-exist if immigrants exiting employment continue job search, while natives leave the labor force after quitting their job. No matter the reason for the observed trajectories in employment outcomes, there is a valid concern that the time-varying unobserved propensity to work might be correlated with earnings profiles, creating the problem of selection bias.

The selection issue is even more concerning when we look at the survey attrition probabilities, also shown in Figure 3. The attrition rate for natives is low and follows a

[8] The same conclusion can be reached from the life-cycle profiles estimated with individual fixed effects. We publish these profiles in appendix Figure W1. Evidence of wage divergence remains strong even after controlling for the immigrant's country of origin, pre-migration background, any factor influencing the past migration decision, and all other components of permanent individual heterogeneity.

normal U-shape trajectory; the annual survey exit rates are about 7 percent at ages 30 and 50 and only 5 percent at age 40. However, the attrition rates for immigrants are considerably higher, with the annual survey exit rate falling between 9 and 19 percent. We do not know the reasons for such high attrition. We can only speculate that, after a temporary stay, many immigrants leave Germany for either their home country or an alternative destination. If immigrants with less favorable prospects in the host country are more likely to leave, the estimated earnings profiles are going to be biased upward. Conversely, positive selection into out-migration (e.g., if high-skill immigrants move away first) would produce the downward bias in assimilation profiles; see Dustmann and Görlach (2015) for an excellent discussion and derivation of biases due to non-random emigration. While our study focuses mainly on wage outcome, we attempt to adjust wage convergence rates for selective out-migration and selective propensity to work by using the Heckman selection correction and inverse propensity weighting procedures discussed in Section 4.

3. Theoretical Model of Wage Convergence

In this section, we present a simple model of wage growth that highlights the main channels behind economic convergence/divergence in wages between immigrants and natives.

3.1 Set-up

The model is based on the standard two-period model of optimal human capital accumulation presented in Borjas (1999, 2015). Borjas's model provides a good starting point in explaining the economic assimilation of immigrants with preexisting human capital, skill transferability, and skill complementarity in the human capital production function. However, as acknowledged by the author, the model always predicts a higher wage growth for immigrants than for natives. Since immigrants are unable to transfer all of their human capital in the host country, they have a lower opportunity cost of investing in human capital than natives. As a result, immigrants invest more in acquiring human capital and experience faster wage growth than comparable natives.

To make the model less restrictive in its predictions and better fit with the empirical facts presented in the previous section, we extend the Borjas model in two major ways. First, we allow the technology of human capital production to differ between migrants and natives. It is quite possible that the lack of institutional/cultural knowledge and language ability could make immigrants less efficient in producing human capital than natives, thus leading to a lower rate of human capital accumulation among immigrants and lower wage growth.

Second, we introduce the price per unit of human capital and allow this price to be different between immigrants and natives.[9] These differences in prices for comparable skills may reflect the lack of information about immigrants' skills (statistical discrimination), distaste, or other forms of labor market discrimination. If prices change differentially between the two groups, this would also affect the optimal amount of investment made in human capital and in turn affect the rate of wage growth.

The rest of the set-up is similar to Borjas (2015). An immigrant arrives to the host country with a stock of pre-migration human capital K, of which τ can be transferred to the host country. Thus, only τK can be used to produce earnings in the labor market. An immigrant lives for two periods in the host country. During the first period, the immigrant decides to invest π of preexisting human capital towards production of new human capital and during the second payoff period he experiences an increase in marketable skills by $g \times 100$ percent. The new human capital ($g \times \tau K$) is produced by investment in the first period (πK) with the use of old human capital K as follows:

$$g\tau K = A(\pi K)^{\alpha} K^{\beta}, \text{ with } \alpha < 1, \tag{2}$$

where A is the human capital technology parameter; α and β are standard elasticity parameters indicating whether new investment and old human capital are substitutable ($\alpha + \beta < 1$) or complementary ($\alpha + \beta > 1$). Thus, the rate of growth of human capital can be conveniently expressed as:

$$g = A\pi^{\alpha}\tau^{-1}K^{\alpha+\beta-1} \tag{3}$$

Individuals choose the optimal π by maximizing the present value of their expected earnings in the two periods. The first period earnings are:

$$p\tau K(1-\pi), \tag{4}$$

where p is the average price per unit of human capital in the first period, with $p = E[p_i]$. In the second period, we allow for the average price to change at the rate of $\dot{p}$. Thus, the second period earnings are given by the following expression:

$$p(1+\dot{p})\tau K(1+g) \tag{5}$$

[9] In the standard model, the market-determined rental rate for an efficiency unit is assumed to be one dollar (Borjas, 1999).

An individual maximizes the present value of earnings over two periods to decide the optimal π:

$$PV = p\tau K(1-\pi) + r[p(1+\dot{p})\tau K(1+g)], \tag{6}$$

where r is the discounting factor. Plugging the expression (3) for g gives the following expression for the present value:

$$PV = p\tau K(1-\pi) + r\big[p(1+\dot{p})\tau K\big(1 + A\pi^{\alpha}\tau^{-1}K^{\alpha+\beta-1}\big)\big]. \tag{7}$$

From the first-order condition, we derive the optimal value of π assuming that the second-order condition holds:

$$\pi^* = \left[\frac{rA\alpha(1+\dot{p})K^{\alpha+\beta-1}}{\tau}\right]^{\frac{1}{1-\alpha}} \tag{8}$$

and,

$$g^* = [r\alpha(1+\dot{p})]^{\frac{\alpha}{1-\alpha}}\left[\frac{AK^{\alpha+\beta-1}}{\tau}\right]^{\frac{1}{1-\alpha}} \tag{9}$$

3.2 Immigrant wage growth

Now let $\dot{w}$ denote the wage growth of immigrants between the two periods. Using equations (4) and (5) with a first order Taylor series approximation, we obtain

$$\dot{w} = log\left(\frac{p(1+\dot{p})\tau K(1+g)}{p\tau K(1-\pi)}\right) \approx g + \pi + \dot{p}. \tag{10}$$

We can express $\dot{w}$ as a function of the endogenous decision variable π,

$$\dot{w} \approx ([r\alpha(1+\dot{p})]^{-1} + 1)\pi + \dot{p}, \tag{10'}$$

or, alternatively, in the reduced form,

$$\dot{w} \approx ([r\alpha(1+\dot{p})]^{-1} + 1)\left[\frac{rA\alpha(1+\dot{p})K^{\alpha+\beta-1}}{\tau}\right]^{\frac{1}{1-\alpha}} + \dot{p}. \tag{10''}$$

The last two equations will be needed later in justifying the estimation strategy. From Equation (10''), we obtain the comparative static derivative of wage growth with respect to parameters of interest.

$$sign\left(\frac{\partial\dot{w}}{\partial K}\right) = sign(\alpha+\beta-1) \tag{11a}$$

$$\frac{\partial \dot{w}}{\partial \tau} < 0 \tag{11b}$$

$$\frac{\partial \dot{w}}{\partial r} > 0 \tag{11c}$$

$$\frac{\partial \dot{w}}{\partial A} > 0 \tag{11d}$$

$$\frac{\partial \dot{w}}{\partial \dot{p}} > 0 \tag{11e}$$

The first three results are described in greater detail in Borjas (2015). Briefly, Equation (11a) shows that high-skill immigrants experience higher wage growth only when their pre-migration human capital is complementary with post-migration investment ($\alpha + \beta > 1$). The next two equations (11b) and (11c) imply that the expected rate of wage growth is larger for immigrants who have lower skill transferability τ and who put higher valuation on future income (i.e., have a bigger r). An example of the latter category of immigrants could be refugees facing a higher cost of return migration. Refugees may appreciate their stay in the host country more and be eager to invest in acquiring new human capital.

The fourth equation (11d) implies that immigrants who are more efficient in human capital production are likely to assimilate faster. This result allows for a more nuanced prediction with respect to some common assimilation factors that appear in both τ and A. For instance, close linguistic proximity of home and host countries may foster higher skill transferability between two countries and lead to slower post-migration wage growth. At the same time, fewer language barriers could make learning new skills more efficient and result in higher wage growth via higher value of A. The same logic applies to ethnic networks. On one hand, a greater number of ethnic compatriots creates a larger market for preexisting skills and reduces incentives to invest into the new set of skills needed in the host country. This corresponds to higher τ and lower $\dot{w}$ in the model. On the other hand, larger, well-established ethnic networks may have institutions in place to make the transition process smoother by providing assistance in acquisition of new skills and thus increasing wage returns per unit of investment (the positive A effect). In both examples, assessing the net effect on wage growth becomes an empirical issue since the theoretical model cannot pin down the direction of the net effect.

Finally, equation (11e) shows that wage changes are responsive to price innovations via both the direct effect of $\dot{p}$ on $\dot{w}$ and the indirect effect of anticipated $\dot{p}$ on investment decisions π^* and subsequent growth in human capital g^*.

3.3 Wage convergence/divergence between immigrants and natives

So far, our theoretical discussion centered on immigrants' own wage growth. While own progress is an important aspect of the economic assimilation, the positive rate of wage growth does not imply that wages of immigrants are necessarily converging to the wage level of natives, as we saw in the previous section. In the notation of our model, the primary statistic of interest here is $(\dot{w}_m - \dot{w}_n)$ or the difference in wage growth between immigrants and natives, where m and n subscripts denote migrants and natives, respectively.

Let us consider a comparable native who is deciding on further investment in human capital and who is identical to an immigrant in terms of the level of pre-investment skills, price per skill, and the technology of human capital production. The wage growth equation for the native is determined by Equation (10''), apart from the skill transferability parameter τ_n, which is equal to 1 since natives can use all preexisting human capital units K in the labor market. In this case, the wage growth differential between immigrants and natives is always positive:

$$sign(\dot{w}_m - \dot{w}_n)|_{A,K,\dot{p},r} = sign\left(\left(\frac{1}{\tau_m}\right)^{\frac{1}{1-\alpha}} - 1\right) > 0, \tau_m < 1 \quad (12)$$

Equation (12) implies that wages of immigrants and natives are converging over time, with immigrants exhibiting higher rate of wage growth in host country, when contrasted to comparable natives. However, relaxing the strict comparability assumption makes the convergence prediction less obvious. Suppose natives have an efficiency edge in human capital production due to better institutional/cultural knowledge, language proficiency, better personal contacts through friends and relatives, and other favoring conditions such that $A_n > A_m$. Then, the wage growth differential between immigrants and natives, even if they have the same starting level of human capital K and face the same price innovations $\dot{p}$, is no longer unambiguously positive. Furthermore, wage divergence becomes possible if $A_n/A_m > 1/\tau_m$:

$$sign(\dot{w}_m - \dot{w}_n)|_{K,\dot{p},r} = sign\left(\left(\frac{1}{\tau_m}\right)^{\frac{1}{1-\alpha}} - \left(\frac{A_n}{A_m}\right)^{\frac{1}{1-\alpha}}\right) \lesseqgtr 0. \quad (13)$$

Similarly, the ambiguity in the wage growth differential emerges when the price dynamics are different between immigrants and natives. A mathematical expression for the relation between $(\dot{w}_m - \dot{w}_n)$ and $(\dot{p}_m - \dot{p}_n)$ is long and complicated, as derived in Appendix A3. But if we fix the price change for one group (e.g., immigrants), then it can be shown that $\partial(\dot{w}_m - \dot{w}_n)/\partial\dot{p}_n < 0$. In other words, the risk of wage divergence is increasing when price innovations favor natives over immigrants.

4. Empirical Strategy

In this section, we discuss the empirical strategy for estimating the model of wage convergence, with a special emphasis on both measuring the factors of economic assimilation and addressing the selectivity issues.

4.1 Empirical model of wage convergence between immigrants and natives

Using individual-level panel data, we build upon the aggregate cohort-level model of wage convergence presented in Borjas (2015). The assimilation rates that Borjas employs as a dependent variable are aggregated from U.S. Census data. They capture the 10-year wage growth experienced by an immigrant cohort from a given country of origin relative to the wage growth experienced by comparably aged native workers. Unlike the cohort-level approach in Borjas's study, our rates of assimilation are individual-specific and vary with individual characteristics at arrival, post-migration investment, and characteristics of home country at the time of entry. Not only we can learn more regarding the sources of individual variation in the rates of wage convergence/divergence over the life-cycle, we can also test several hypotheses, including the role of post-migration investment in wage divergence, which the cohort-level analysis cannot do.

If we ignore the issues of endogeneity and selectivity for a moment, the individual-level model of wage convergence can be expressed in a single linear equation:

$$\dot{\theta}_{it} = \beta_0 + \beta_F F_{it} + \beta_\varphi \varphi(AGE_{it}, TIME_{it}) + u_{it}, \quad (14)$$

The dependent variable in this model is the average annual change in relative wages over the next 5-year period, $\dot{\theta}_{it} = (1/5)\sum_t^{t+4}(\theta_{it+1} - \theta_{it})$ for immigrant i at time t. We had to make several choices in constructing the dependent variable. The wage measure is hourly, and it is calculated as the total net income earned from employment last month in constant 2010 euros divided by the product of actual working hours per week and the number of weeks in a month. Actual hours are chosen over contractual hours because actual hours are available for the self-employed and include over-time work. Between net income and gross income, we choose the former as individual work and migration decisions are influenced by the net income. We only use non-imputed income because earnings imputation can cause match bias, as shown by Bollinger and Hirsch (2006). Observations with income imputed by the GSOEP are treated as missing and modeled as part of the selection process.

The change in relative wages is calculated over the 5-year period. The 5-year interval is not too short to be overly sensitive to transitory earnings shocks and measurement error. On the other hand, it is not too long to lose a significant number of observations due to survey attrition and outmigration. In calculating the 5-year average rate of relative wage growth, we use the minimum of three non-missing data points. This allows to retain immigrants who temporarily drop out of employment or leave the survey for 1 or 2 years.

Our preferred measure of the dependent variable is the change in conditional relative wage of immigrants. Recall from Section 2 that the conditional relative wage shows the immigrant's position in the wage distribution of comparable natives of the same age, schooling, and location type. However, we also use the absolute wage growth of immigrants as a dependent variable.

One of the advantages of the individual-level wage difference model is that it differences out permanent unobserved individual heterogeneity in the level equation, including all characteristics of the immigrant at the time of arrival such as entry wage, age-at-migration, unobserved skills, location at arrival, family background, pre-migration history unknown to the econometrician, among others.[10]

The covariates that influence the trajectory of relative wage in Equation (14) include a flexible function of the immigrant age and survey time, $\varphi(AGE_{it}, TIME_t)$, as well as the vector of other observed factors of wage convergence, F_{it}, which comprises of individual characteristics at the time of arrival, such as gender, ethnicity, and parents' education; time-varying individual characteristics in the host country, including post-migration investment in human capital and location in Germany; characteristics of the home country at the time of arrival (e.g., linguistic proximity, GDP per capita, and political violence); time-varying home country variables such as the size of ethnic networks; and time-varying destination country characteristics such as economic growth in Germany. Next, we discuss the rationale for why each variable is chosen.

[10] The first difference estimator has been previously used in the immigration literature with respect to more aggregated units of analysis, such as cities (Altonji and Card, 1991), skill groups (Dustmann *et al*., 2010), or arrival cohorts (Borjas, 2015).

4.2 Measuring factors of wage convergence

Our choice of convergence factors entering the vector F_{it} is guided mainly by the theoretical model of wage convergence presented in Section 3. The model distinguishes between the endogenous choice variable π indicating post-migration accumulation of human capital and the set of exogenous factors $(K, r, \tau, A, \dot{p})$ influencing the wage growth of immigrants directly or indirectly through π; as shown by Equations (10') and (10'').

We begin with measuring π and K. The importance of splitting human capital into pre- and post-migration components has long been recognized in the immigration literature focusing on the wage returns to human capital (Bratsberg and Ragan, 2002; Chiswick and Miller, 1994; Ferrer *et al.*, 2006; Sanroma *et al.*, 2015; Skuterud and Su, 2012). Our data allow us to separate education acquired in the home country from post-migration investment in the host country. We do it by using the age-at-migration and annual spells of schooling and job training between the ages of 15 and 65 (see Appendix A2).[11] In measuring preexisting human capital K at the time of arrival, we use not only years of education acquired in the home country, but also the highest level of schooling completed by a parent. These variables are predicted to have either positive or negative impact on immigrants' wage convergence depending on whether K is complementary or substitutable with post-migration investment, as in Equation (11a). In measuring π, we observe whether the immigrant studied in the German school and/or underwent job training after migration; both factors are predicted to have the positive effect on wage convergence according to Equation (10').

The third factor is the valuation of future income (r), which in the model leads to more human capital accumulation and higher rate of wage growth after the arrival. Although there is no direct measure of this factor, previous studies suggest that immigrants who escaped political instability and violence in their home country may place a higher value on their future in the new country, have a lower likelihood of returning home, and hence invest more in the host country (Borjas, 2015; Chin and Cortes, 2015; Cortes, 2004). Following this line of argument, we use the annual index of political instability in home country at the time of arrival to differentiate between

[11] It is common in the immigration literature to calculate years of schooling in the host country as total years of schooling plus 6 or 7 years of school starting age minus age of migration, thus assuming that schooling is continuous and is not interrupted by the transition from one country to another (Bratsberg and Ragan, 2002; Friedberg, 2000; Sanroma *et al.*, 2015). This procedure tends to underestimate years of schooling completed in the host country. Having detailed spell data before and after migration in the GSOEP alleviates potential measurement error, but with one caveat. We had to award each immigrant with equal years of schooling before age 15 (7 years at age 14, 6 years at age 13, and so forth). Since our focus is on adults migrated at age 15 or older, adding this constant should not have an impact on estimated slopes.

immigrants with different values of r. The index is published by the Center for Systemic Peace (2015). It assesses major episodes of international, civil, and ethnic violence and warfare for almost 180 countries worldwide between 1946 and 2014. Based on the index of political instability, we split all country-year observations into four categories: no episodes of political violence, limited political violence, serious political violence, and warfare.

The fourth assimilation factor is the level of skill-transferability (τ), which is shown to be inversely related to the rate of economic assimilation.[12] Borjas (2015) posits that the skills of immigrants are more easily transferable between the two industrialized economies. We use the log of real GDP per capita in the home country at the time of arrival to capture the level of skill transferability. The underlying expectation is that immigrants from a low-income country would have to invest more into the skills relevant to the advanced host country and experience larger wage gains through acquiring new skills and information. Other studies propose using linguistic distance/proximity between home- and host-country languages as a measure of skill transferability (Chiswick and Miller, 2012). Indeed, immigrants who grew up speaking the language that is distant from German face higher cost in the transfer of their preexisting skills to the new labor market. A lower value of τ implies a steeper earnings profile, according to Equation (11b). The measure of linguistic proximity is described in Appendix A2.[13] While capturing skill transferability, both GDP per capita and linguistic proximity may also depict the efficiency differences in the production of human capital. For example, linguistic barriers could make immigrants less efficient in learning new skills (i.e., have lower A), thus slowing subsequent wage growth. Lower levels of economic development in the home country could be associated with poor school quality and inadequate learning practices, which may hinder the effectiveness of new skill acquisition in the host country. Empirically, we can test which channel (via τ or A) dominates.

A similar ambiguity arises with respect to the role of ethnic networks, which are often measured as the share of total population from the same country of origin in each geographic area or in a host country at large. Borjas (2015) argues that larger ethnic

---

[12] Similar prediction is made by Duleep and Regets (1999) who theoretically rationalize that immigrants with less-transferable skills would start at a lower level of earnings, but experience faster earnings growth due to both greater human capital investment and higher value of host-country skills.

[13] The measure of linguistic proximity is based on the *Ethnologue* classification of language family trees (Ethnologue, 2016). It takes five ordered values between 0 (two languages belong to different trees) and 1 (German is an official language) depending on how far primary home-country language is from Standard German in the linguistic tree. In the immigration literature, a similar measure has been used by Adsera and Pytlikova (2015).

networks, by creating demand for preexisting skills and improving skill transferability, reduce incentives for learning new skills and discourage socio-economic exchanges with natives. He documents a significant negative effect of the number of preexisting immigrants from a given country on a 10-year wage growth for new arrivals from the same country. This result is consistent with the skill transferability channel. However, the positive assimilation effect of networks via the efficiency channel is also plausible. Existing networks could be helpful to newcomers in acquiring new skills through, for example, on-the-job training or shared information about training opportunities. The access to formal education is likely to be less costly in the areas with high concentration of immigrants due to scale economies. Therefore, the effect of ethnic networks on human capital accumulation and subsequent wage growth is ambiguous.[14] In testing which channel dominates, we include the relative size of ethnic groups at the beginning of the 5-year period (and alternatively at the time of arrival) as a covariate in the wage growth model. The size is calculated as the share of foreign population by country of origin in the total German population.[15]

The effect of ethnic networks on the economic assimilation of immigrants can vary depending on how old the networks are. Well-established networks may have institutions in place for better social and economic integration of immigrants. Munshi (2003) shows that individuals belonging to established networks have a higher probability of employment than immigrants from relatively new networks. Hatton and Leigh (2011) also find that immigrants from countries with long history of migration gain relatively more, however, the geographic concentration of ethnic networks depresses earnings of immigrants. We check whether the age of networks matters for wage assimilation by splitting our network measure into (i) established networks based on the stock of immigrants from the same country of origin 5 years ago and (ii) recent networks measured as additional flows of immigrants from the same country of origin during the last 5 years. The average share of ethnic groups in the foreign-born population of

---

[14] Several previous studies find that living in an ethnic enclave improves earnings of immigrants (Damm, 2009; Edin *et al.*, 2003). Since earnings in these studies are measured in levels, not in differences, these findings would be consistent with either transferability or efficiency channel and with either negative or positive wage growth. Higher levels of initial earnings resulting from the network effect may imply better skill transferability and lead to lower earnings growth. Conversely, observed earnings gains associated with living in ethnic enclaves could be an outcome of better resource allocation in the production of new skills (i.e., higher A) and be associated with higher earnings growth.

[15] Ideally, the size of ethnic networks should be adjusted for the degree of regional clustering of immigrants, as in Borjas (2015). We do not have such information for Germany. The scarce evidence suggests that the geographic segregation of ethnic groups in Germany is relatively low and stable over time (Constant *et al.*, 2013).

Germany has been on decline since the 1970s. We find that the Herfindahl index of ethnic concentration fell considerably from 0.23 in the 1970s to 0.19 in the 1990s, 0.12 in the 2000s, and less than 0.07 in 2015, thereby indicating increasing diversity of foreign-born population in terms of origins.[16]

Individual-level data allow us to look at other sources of individual variation in wage convergence rates. Our theoretical analysis is constrained to a stylized setting with limited factors, but empirically we can test for differences in convergence rates by gender, age, urban residence, and German ethnicity. The ethnicity factor could be linked to the theoretical model, as it reflects both skill transferability and the efficiency of new skill acquisition since ethnic Germans are likely to be more familiar with language and social institutions in Germany.

The last but not least factor affecting economic assimilation of immigrants in our model is changes in the prices of skills. The fundamental problem here is that skill prices are not directly observed in the data. Therefore, price innovations favoring one group (e.g., natives) over the other, after controlling for observed characteristics, are an omitted variable in Equation (14). We estimate Equation (14) under the assumption that post-migration price changes in the host country are uncorrelated with pre-migration individual and home country characteristics and that the price dynamics are part of the residual. In Section 5.3, the native-immigrant wage gap will be related to an available measure of perceived discrimination, which we use as a proxy for the price differential.

4.3 Selectivity bias and exclusion restrictions

When estimating the model of wage convergence as in Equation (14), the problem of selection bias arises from missing data on the dependent variable. For simplicity of presentation, let suppress subscripts i and t and combine all exogenous covariates in Equation (14) into one vector X:

$$\dot{\theta} = \beta X + u, \tag{15}$$

This empirical equation corresponds to the reduced-form theoretical Equation (10''), as it replaces the endogenous post-migration investment in human capital with the set of exogenous assimilation factors. We assume that dropped observations with missing X's are ignorable or missing completely at random since the share of missing values in the X

[16] We publish these and other trends in macro indicators of home countries in web appendix Figure W2, where readers can see how political instability, GDP per capita, linguistic proximity, ethnic networks evolve over time. These home-country characteristics display substantial variation both over time and across countries of origin.

vector is less than 5 percent of the raw data. However, the issue of non-random sample selection due to missing dependent variable $\dot{\theta}$ cannot be ignored, particularly when the dependent variable represents the growth or change. Growth variables may lead to a greater loss of valid observations in the estimation sample, which could be related to out-migration, respondent's death, non-response at follow-up, exit from employment between the two survey rounds, or simple refusal to report a wage. In addition to traditional "static" selectivity bias due to non-participation in the labor market at any given time, the non-random survey attrition and employment exit between survey rounds, if correlated with u, can cause the OLS estimation of relative wage growth equation to be biased and inconsistent.

We use three alternative strategies to account for selectivity bias: inverse propensity weighting (IPW), the Heckman 2-step procedure with the inverse Mills ratio (IMR), and full maximum likelihood estimation (MLE). All three approaches are based on the same selection equation. Let D be a selection dummy variable, taking the value of one if $\dot{\theta}$ is observed and the value of zero if the dependent variable is missing. Here, we do not distinguish between different reasons for missing data on the dependent variable.[17] The selection dummy is assumed to be linked through the indicator function $D = I(D^* > 0)$ with the following latent index model:

$$D^* = \gamma Z + \upsilon, \tag{16}$$

where Z is a vector of explanatory variables that are observed for all individuals; v is an *i.i.d.* error term with mean zero and unit variance. The vector Z includes all variables from the vector X as well as exclusion variables Z_1 affecting the selection process, but not included in the wage growth equation. Under standard exogeneity assumptions, $cov(Z_1, \upsilon) = cov(Z_1, u) = 0$. Specifically, we use *characteristics of the interview* to predict survey attrition. In the GSOEP, we know the mode of interview in year t, such as face-to-face, computer assisted, and self-written and mailed. We also know whether it was a first-time interview and for subsequent interviews whether the same interviewer surveyed the household in year t as in year t-1. Interview characteristics are generally found to be good predictors of continued survey participation. We also include *the average annual growth of real GDP per capita in the home country* during the next 5-year period [t+1, t+5] as a potential factor influencing the decision of immigrants to stay in the host country. The

[17] Initially, we jointly estimated two selection equations for survey attrition and employment participation. However, since the splitting of the selection model into two sub-models made no difference for our final results, we choose a simpler one-equation selection model. The two-equation selection model is shown in Appendix Table A3-2.

worse the economic conditions in the home country, the higher the cost of returning home and the lower the expected probability of out-migration (and thus the likelihood of survey drop-out in the next year). Finally, we use *the average commuting distance between home and workplace* by state-year. This variable serves as a proxy for fixed costs associated with work, and it is likely to influence the employment participation decision. We note that the average commuting distance is calculated for all workers, including natives, and it is one-year lagged relative to future 5-year wage growth of immigrants. This is done in the attempt to mitigate potential feedback effects of the immigrants' wage growth on job density in the local labor market.

Once Equation (16) is estimated using the probit model, the IPW procedure applies the weighted least square method to Equation (15), where the weights are given by the inverse of $\Phi(Z\hat{\gamma})$; see Rosenbaum and Rubin (1983), Hirano *et al.* (2003), Wooldridge (2007). The Heckman 2-step selection correction includes the inverse mills ratio, $\lambda = \phi(Z\hat{\gamma})/\Phi(Z\hat{\gamma})$, in the estimation of Equation (15) as an additional generated regressor. The MLE method jointly estimates Equations (15) and (16) under the assumption that the error terms in two equations (u, v) are jointly normally distributed. The joint normality assumption is also required for the Heckman 2-step procedure, but it is not necessary for the IPW method. However, the IPW method maintains the unconfoundness or "selection on observables" assumption, while the other two methods allow for "selection on unobservables", which is generally preferred.

4.4 Endogeneity of post-migration investment in human capital

The theoretical model of wage convergence makes it apparent that post-migration investment is endogenous. The immigration literature has not yet addressed this issue, except for adding individual fixed effects in the wage-level equation (Skuterud and Su, 2012). In our case, the differencing of relative wage removes the permanent component of individual heterogeneity. In addition, we use different time periods for two processes to avoid immediate simultaneity: post-migration education is acquired between a (arrival year) and t (current year), whereas the wage growth is observed during the next 5-year period [t+1, t+5]. However, lagged investment in education could still be endogenous in the growth/change equation since individuals may base their decisions about new skill acquisition not only on anticipated wage levels and contemporaneous wage growth but also on the entire wage trajectory, which includes future wage growth.

As we saw in Table 1, only 14 percent of immigrants receives formal schooling in Germany and 14 percent acquire post-migration job training (5 percent has both). Most immigrants who study in Germany attend short-term programs for less than 2 years. Given that the distribution of years spent on investment in post-migration human capital is highly skewed, we dichotomize years of education into a binary variable. Thus, our wage convergence model can be characterized as a standard linear regression model with an endogenous binary treatment variable (T).

We use two econometric methods to estimate the treatment effect of post-migration education on immigrants' wage convergence. The first method is a doubly robust average treatment effect (ATE) estimator described in Wooldridge (2010). The estimator is known as IPWRA because it combines regression adjustment with inverse propensity score weighting to provide double robustness to the estimate. The estimator models the selection-into-treatment equation (17) and two potential outcome equations (18) for T=0 and T=1 as follows.

$$T = \begin{cases} 1 \text{ if } \alpha W + \eta > 0 \\ 0 \text{ otherwise} \end{cases} \tag{17}$$

$$\begin{aligned} \dot{\theta} &= T\dot{\theta}_1 + (1-T)\dot{\theta}_0 \\ \dot{\theta}_0 &= \beta_0 X + u_0 \\ \dot{\theta}_1 &= \beta_1 X + u_1 \end{aligned} \tag{18}$$

The ATE is estimated in three steps. The first step involves estimating the probit treatment model (17) and calculating the propensity score $\hat{p}(W_i, \hat{\alpha})$. The second step estimates two outcome models jointly by quasi-maximum likelihood using inverse propensity weights, $1/\hat{p}(W_i, \hat{\alpha})$ for the treated group and $1/(1-\hat{p}(W_i, \hat{\alpha}))$ for the control group. The final step calculates the ATE as the average difference of predicted values:

$$\Lambda_{ATE} = N^{-1} \sum_{i=1}^{N} [(X_i\hat{\beta}_1) - (X_i\hat{\beta}_0)]$$

The IPWRA 3-step estimator provides a consistent estimate of ATE even if one of the models is misspecified. In other words, we only need either the conditional mean outcome model or the propensity score model to be correctly specified. The estimator requires each observation in the sample to have a positive probability of treatment, which is known as overlap assumption. We use the propensity score $\hat{p}(W_i, \hat{\alpha})$ in the range [0.001, 0.999] to ensure the overlap assumption is not violated. The estimator also requires conditional mean independence (CMI). That is, after conditioning on covariates, the treatment assignment should not affect the conditional mean of each potential outcome. To the best of our efforts, we control for the large number of factors that could affect post-migration education. In addition to ATE, we also estimate the average treatment effect on the treated (ATT), which relies on a weaker CMI assumption: unobservables in the treatment model should be uncorrelated with the conditional mean outcome of the control group only (Wooldridge, 2010).

The second econometric approach attempts to relax the CMI assumption by using the standard instrumental variable approach for estimating the local average treatment effect (IV-LATE). We employ demand-supply shifters in government-sponsored training programs as IVs. Specifically, using annual statistics on the demand and supply of training contracts for West Germany, we construct the log of supply of training offers and the excess demand for training at the arrival a and use them in the equation for wage growth during [t+1, t+5]. The excess demand is calculated as the difference between the log of unplaced training applicants and the log of unfilled training places still registered with employment offices. Figure 4 shows significant fluctuations in these measures over time. We assume that the supply of training offers and the excess demand at arrival influence new skill acquisition in the host country between a and t, but they are not forecastable based on future (residual) long-term wage growth between t+1 and t+5.[18]

The shortcoming of demand-supply shifters is the lack of individual variation. They can only be used in conjunction with some form of a polynomial in time trend but not with year fixed effects. We introduce individual variation by using another IV, which measures potential interruptions in early schooling due to wars and political violence in the home country. Specifically, we calculate the average political violence score in the country of origin during the time when an individual was 6 to 10 years old. Interruptions in elementary schooling due to wars and internal conflicts are likely to have an adverse effect on subsequent schooling trajectory and thus affect the accumulation of human capital after migration. We assume that this IV is uncorrelated with unobservables in the wage convergence equation. If the dependent variable was measured in levels, this assumption would be problematic since conflicts may affect immigrant's permanent cognitive and non-cognitive ability. In our case, the permanent unobserved ability is differenced out. All we need to assume is that the change in unobserved ability after migration is independent of the episodes of political violence experienced during early schooling age.

There are known problems with the consistency of 2SLS with dummy endogenous regressors; see Angrist (2001). To account for the endogeneity of the binary variable, we

[18] Given the substantial number of unfilled vacancies and unplaced applicants, the government is not a good forecaster even over the one-year horizon, not to mention longer periods sometime in the distant future. The series is only available from 1976 and onwards and we had to assign the 1976 value to earlier years.

implement the 3-step procedure (Probit-2SLS) recommended by Wooldridge (2010). In the first step, we obtain predicted probabilities of studying in the host country π, $G(X, IV; \gamma_S) = Pr(\pi = 1 \mid X, IV)$, where X includes the same variables as in Equation (15); G (.) is a probit model. Then, we estimate 2-stage structural Equation (19) by instrumenting endogenous π with IVs and nonlinear fitted values $\hat{G}$.[19]

$$\dot{\theta}_{it} = \beta_0 + \beta_\pi \pi_{it} + \beta_X X_{it} + u_{it} \tag{19}$$

As with any IV approach, there is a valid concern that the local ATE it captures may be far off from the true ATE. The IVs themselves could be fairly criticized for a likely violation of exogeneity assumption. However, by utilizing different approaches and methods, we hope to get a consistent picture for the overall effect of post-migration investment in human capital on the immigrants' wage convergence and to test the efficiency channel for the observed wage divergence.

5. Results

5.1 Reduced form model of wage convergence

We begin our data analysis by providing summary statistics and presenting the estimates of the reduced-form specification of wage convergence. Table 3 shows the mean and standard deviation of model covariates for the two samples: (i) a censored sample with non-missing $\dot{\theta}_{it}$ and (ii) a combined censored and uncensored sample used in estimating the probit selection model. Because different covariates are measured at different points in time, we draw the time frame below to help to visualize the timing of covariates. This picture illustrates the data structure for a hypothetical immigrant who arrived in Germany in 1994 and is interviewed in year 2000. All background variables, home country characteristics, and pre-migration investment in human capital are measured at the time of arrival, a. Post-migration schooling and training, current location in Germany, interview characteristics, and the average commuting distance are taken at the year of the interview, t. Finally, the wage growth and economic growth are calculated as a moving average over the future 5-year period [t+1, t+5]. Thus, in our estimation sample, the year 2009 is the last possible year of migration for which the growth rates can be constructed. The details on each variable are provided in Appendix A2.

[19] Equation (19) may include the predicted inverse mills ratio to deal with non-random sample selection, in which case standard errors need to be bootstrapped.

Time Frame

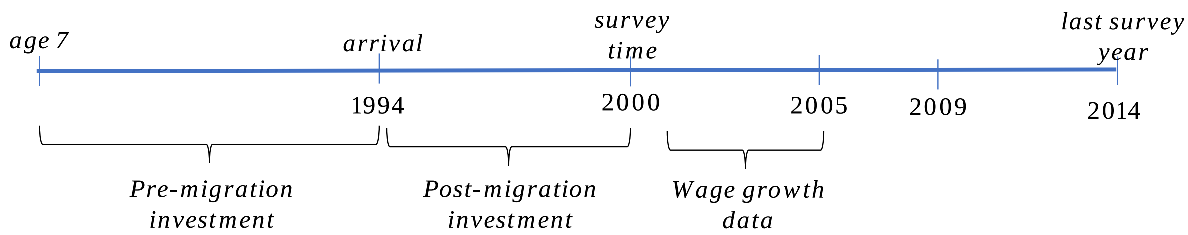


Table 4 reports the baseline model estimates of Equation (15) for the change in conditional relative wage using different methods of selectivity bias correction. Each column represents one of the four specifications – OLS, IPW, IMR, and MLE. The results are mostly consistent across specifications, with only two notable exceptions. First, the OLS, IMR, and MLE approaches generate a statistically significant positive effect of pre-migration education on wage convergence, while the IPW method shows no such effect. Second, the IMR method does not find statistically significant gender differences in wage convergence, while the other three methods estimate higher wage convergence for female immigrants.

We find that immigrants with college-educated parents are more likely to experience wage convergence. Their gap with comparable natives is closing by almost 1 percentile per year. The estimated positive coefficient on both pre-migration education and investment by parents may be indicative of the complementarity between pre- and post-investment in human capital, according to Equation (11a) in the theoretical section. If we split pre-migration education into formal schooling and job training, as we do it in Panel B of Table 7, the positive effect of preexisting human capital on wage convergence appears to be largely driven by home-country formal schooling. The job training received in the home country does not seem to influence the immigrant's rate of wage progression relative to natives.

Only one measure of skill transferability – the linguistic proximity between home- and host-country languages – is estimated to have the negative effect on the relative wage growth, as predicted by Equation (11b). Closer proximity may weaken the incentives of immigrants to acquire new skills in the host country. Yet, other variables that have been traditionally used to measure skill transferability, such as GDP per capita in the home country, German ethnicity, and the size of ethnic networks, are all estimated to have the positive effect on wage convergence. Although these results are not consistent with the skill transferability explanation alone, they make sense according to the efficiency differences in human capital production function. Higher GDP per capita in the home country might be depicting better school quality or familiarity with educational institutions comparable to Germany, which can make the acquisition of new skills and job search easier. Ethnic German resettlers, who are likely to be better culturally assimilated, may have an edge in acquiring host-country skills. This would explain the faster progression of their wages to the level of natives, as compared to other ethnicities by about 0.3 percentiles per year.

Likewise, larger networks may help immigrants in the assimilation process by disseminating valuable information, providing contacts, and finding jobs with better growth prospects. If we split ethnic networks into established and more recent networks (measured as the population share of compatriots 5 years ago and the share of new arrivals from the same country of origin during the last 5 years, respectively), we find that only established networks have a positive and statistically significant effect on wage convergence, and no such effect is found for more recent networks; see Panel C of Table 7. Several other studies also find the positive role of well-established networks in the economic assimilation of immigrants (Munshi, 2003; Hatton and Leigh, 2011).

Another interesting observation comes from the age-related effects. As we saw earlier in Figure 2, the wage divergence between natives and immigrants (or the slope difference) is highest during early adulthood, which is the period of human capital accumulation. This conclusion holds after controlling for other assimilation factors, as shown in Table 4. Immigrants in the youngest age group (17-25) exhibit the lowest rate of wage convergence.

The relationship between political instability and wage convergence is positive, which agrees with the prediction of our theoretical model. Immigrants from warfare zones, who are likely to be refugees and asylum seekers, assimilate faster. Generally, refugees do not have an option of going back to their home country and value the future stream of earnings in the host country. Their faster wage convergence could be related to stronger incentives to invest more in host-country-specific human capital.

We also observe some spatial and temporal dispersion in wage convergence. Specifically, immigrants in rural areas have a faster rate of wage assimilation relative to immigrants in urban areas. The estimated coefficients associated with year fixed effects that are depicted in Figure 5 exhibit significant fluctuations over time. These fluctuations do not appear to follow the business cycle, as we observe in Panel A of Table 7, where year fixed effects are replaced with a linear time trend and the average annual growth of per capita GDP in Germany over the 5-year period. At the same time, the secular trend in wage convergence is tilted slightly upward, as can be seen in both Figure 5 and Table 7A. This finding is different from the U.S. where the rates of economic assimilation have been declining (Borjas, 2015). It could be that the convergence slowdown depicted by Borjas is U.S. specific and not generalizable to other developed countries.

Lastly, we notice that selection terms (inverse Mills ratio in IMR and the correlation between the error terms in MLE) are not statistically significant. The divergence of relative wages does not appear to be attributed to the selection effects associated with survey attrition, out-migration, selection into employment, non-response, and other sources of missing wage growth. Selection terms and inverse propensity weights have been constructed using the selection equation given by Equation (16). The estimates from the selection equation can be found in Table 5. All exclusion restrictions are significant predictors of the selection into observed wage growth. The selection probability is lower for the first-time survey participants and when the interviewer has changed from the previous round. Similarly, higher future growth of per capita GDP in the home country decreases the likelihood of observing change in wage in the host country. This result is most likely due to out-migration in response to better prospects in the home country. Expectedly, an increase in the average commuting distance to workplace significantly reduces the probability of employment and non-missing wage growth. Apart from exclusion restrictions, we find that the probability of observed wage growth is increasing with years of education in the host country and linguistic proximity, which is expected. The selection probability is much higher for males, ethnic German resettlers, and middle-aged groups. All these results are intuitive. The only unexpected result is a lower probability of being in the estimation sample for the immigrants who escaped political violence. Further analysis reveals that this last result is driven mostly by the lower likelihood of wage reporting by the immigrants from politically unstable countries rather than their survey attrition or employment exit. This could be related to higher dependency of refugees on informal employment (Tumen, 2016).

In Table 6, the same four methods (OLS, IPW, IMR, and MLE) are applied to a different dependent variable – the average annual wage growth of immigrants over the 5-year period. This dependent variable depicts own progress of immigrants over their life-cycle irrespective of the performance of comparable natives in the labor market. Several interesting observations emerge from these estimates. Consistent with the concave age-wage profiles (also shown in Figure 2), immigrants' wage growth decreases with age. The negative effect of age on own wage progression is expected, but it is opposite to the positive effect of age on wage convergence when we compare immigrants with comparably-aged and -skilled natives. Another difference between the two models involves the effect of pre-migration education, which is estimated to be positive in the wage convergence model, but it is small and statistically insignificant in the wage growth model. This implies that each additional year of home-country education helps to close the wage gap with natives within the same education level but not across education levels. The two models also appear to be different with respect to changes over calendar time, as reported in Panel A of Table 7. The wage growth of immigrants is highly procyclical, but the wage convergence is not. While the average rate of wage growth declines over the 25-year period, the rate of wage convergence exhibits a slight increase. Aside from these differences, all other conclusions are similar between the two models.

5.2 Post-migration accumulation of human capital

Recall from our theoretical model that wage divergence can be generated by the differences in the efficiency of human capital production between immigrants and natives. This prediction also implies that individual variation in efficiency (or A parameter in the production function) among immigrants may lead to different trajectories of wage assimilation. Empirically, we can test if immigrants with a likely advantage in the production of host-country-specific human capital indeed enjoy higher future gains associated with post-migration education. But, first, we need to check whether the new skill acquisition by immigrants in the host country closes the wage gap with natives.

In the first set of estimates presented in Panel A of Table 8, we calculate the ATE and ATT associated with post-migration education of immigrants. Treatment effects are obtained using the IPWRA procedure which stands for the doubly robust regression adjustment with inverse probability weighting. The procedure follows Wooldridge (2010), and it is described in Section 4.4. Alternative estimates using simple IPW and augmented IPW are also shown. We report the auxiliary equations used in calculating the IPWRA treatments effects in Appendix Table W3. These equations consist of one probit equation for post-migration education and the two OLS equations for wage convergence, which are estimated separately for the treatment and control groups. The outcome equations have the same set of covariates as in Table 4.[20] The IPWRA method offers

[20] Given the statistically insignificant effect of sample selection in previous tables, we do not include the selection term in Tables 8 and 9. Adding the inverse Mills ratio into Equations (17) and (18) does not make any difference to the mean treatment effect, but the bootstrapping of standard errors from 1000 iterations would increase the computation time to several days.

flexibility by allowing different covariates in the treatment assignment and outcome models. For example, it makes sense to measure time-varying covariates such as age and ethnic networks at the beginning of the investment period in the selection-into-treatment model, but at the end of the treatment (or right before the growth period) in the outcome model. Although the doubly robust method does not require exclusion restrictions, we add IVs in the probit model since we have them. In IPWRA1, the selection-into-treatment model includes the supply of training offers at arrival, the excess demand for training at arrival, potential schooling interruptions due to political violence from age 6 to 10, and the quadratic function of time. For robustness checks, we also present estimates IPWRA2, which are based on the probit model without IVs but with year fixed effects instead of a time trend. In all estimated specifications, we find positive and statistically significant gains in relative wage growth associated with post-migration investment in education. In our preferred specification, IPWRA1, immigrants who study in Germany close the wage gap with natives by 0.5 percentile per year.[21]

Next, in Panel B of Table 8, we contrast the treatment effect between the two groups of immigrants. The groups are created based on available proxies for the efficiency in human capital production. The first group includes immigrants who are closer to natives in certain characteristics and may have an efficiency edge in new skill acquisition due to linguistic, cultural, or spatial proximity and/or similar levels of development of the home country. For example, ethnic Germans, immigrants from countries whose language belongs to the same family tree as German language, immigrants from neighboring and more developed countries belong to this group. The estimates in Table 8 indicate that post-migration investment in human capital makes these categories of immigrants more successful in catching-up with natives compared to immigrants who do not belong to these groups.

In the last set of results presented in Table 9, we show the IV-LATE estimates of the effect of post-migration education on wage convergence. The results are based on three IV estimators with an additional probit equation: two-stage least squares (2SLS), generalized method of moments (GMM), and limited-information maximum likelihood (LIML). The 3-step procedure is described in Section 4.4. We only report the estimated coefficient on post-migration education variable. Full results from the OLS and IV-LATE estimation are shown in Appendix Table W4. According to the OLS estimates, immigrants

[21] About 200 observations with the propensity score outside the range [0.001, 0.999] are dropped to satisfy the overlap conditions. The standard overlap figure shows a practically full-range overlap in the predicted propensity score between the treatment and control groups (see Figure W3).

who study in Germany reduce their wage gap with natives by 0.28 percentiles per year of staying in the host country, ceteris paribus. The annualized returns in convergence are 0.1 percentile for each additional year of study (=0.28 divided by 2.8 years of average duration of education for the treated). The OLS estimate is on the lower side, about half of the IPWRA estimate, but it is still twice as high as the effect of one year of study in the country of origin. All other estimated coefficients are very close to the ones reported in Table 4.

In the first stage of Probit-2SLS, each IV is statistically significant and has the expected sign. The probability of post-migration education increases with a greater supply of training offers and higher excess demand, and it decreases with higher average political violence during early schooling age in the country of origin. The chi-squared test for excluded instruments is 147.24, significant at the 1-percent level. The IV estimates of the treatment effect are statistically significant and have the expected positive sign. Probit-2SLS does not allow for the standard test of overidentifying restrictions to be implemented since the instrument is the fitted propensity score. However, the magnitude of the estimated coefficients seems implausibly high – 3 percentiles in gap closing per year of stay in Germany. It could be that our instruments are not exogenous, or it could be that our IV estimates capture a very high local ATE for the compliers who respond to IVs.

Despite these concerns with IV estimates, the remarkable result is that very different groups of estimators – OLS, IPWRA and its modifications, and IV and its modifications – produce the same pattern when we split the sample of immigrants based on the proximity of their characteristics to natives. Immigrants who are closer to natives in terms of language, ethnicity, geographic distance, and the level of economic development of their home countries enjoy much greater benefits from the post-migration investment in host-country education compared to other groups of immigrants.

5.3 Channels of wage divergence

As indicated by the theoretical model, there could be at least two channels of wage divergence: (i) the relative efficiency differences between immigrants and natives in the production of human capital (A_n/A_m), and (ii) the change in skill prices favoring natives ($\dot{p}_n - \dot{p}_m$). Although we do not have a direct measure of the technology parameter in the production function, several previous findings are consistent with the first channel. Both Figure 2 and Table 4 show that the immigrant-native wage gap widens the most during the period of young adulthood, which is also the period of the largest investment in post-secondary human capital. Of all age groups, immigrants in the youngest age group (17-25) exhibit the highest rate of wage divergence. The second indirect evidence in support of the efficiency channel comes from the heterogeneous convergence effect of post-migration education in the host country. As we saw in Section 5.2, only select groups of immigrants are closing the wage gap with natives as a result of post-migration education in Germany. These groups are likely to have advantages in the production of host-country-specific human capital (or have a high A) because of language, cultural closeness, spatial proximity, and similar educational and other institutions in the country of origin. For many other groups of immigrants with a likely low A, the wage gains from host-country education are insignificant or not large enough to improve the position of these groups in the wage distribution of comparable natives.

The second channel is more difficult to test. The divergence through the price channel occurs not because of static discrimination when the same skills of different population groups are awarded differently by the labor market, but because of intertemporal discrimination when the market price of the same skills changes differentially with age, favoring one group over the other. An additional challenge is that neither static nor intertemporal discrimination is directly observed. Attributing the entire residual from the wage convergence equation to the discrimination factor would require overly strong assumptions, which we are not comfortable making. The best age-varying measure we have of discrimination against immigrants is perceived discrimination. It comes from the survey question "How often have you personally felt disadvantaged in Germany in the last two years because of your origins?" This question was answered by immigrants in 1996-2011 and 2013 surveys.[22] In our estimation sample of immigrants, 39 percent experience disadvantages sometimes and 7 percent encounter discrimination frequently.

Admittedly, the survey question on perceived discrimination sounds vague, and the meaning of disadvantage is uncertain. Despite such ambiguity, the discrimination variable appears to be consistent with the actual placement of immigrants in the wage distribution of natives. For example, in 2011, immigrants who say they have frequently experienced discrimination place at the 35th percentile of conditional native wage distribution, whereas those immigrants, who never experienced disadvantages, place at

[22] We had to exclude a newly drawn sample of immigrants in 2013 (Sample M) because the formulation of discrimination questions for this sample applies to the previous experience in general and does not refer to the last two years or any other specific period.

the 40th percentile. Without making any causal inferences, in a simple regression of the conditional relative wage on the discrimination variable, age and year fixed effects, we also find that frequent episodes of discrimination are associated with a 5-percentile lower placement in the native wage distribution compared to no episodes of discrimination; immigrants experiencing disadvantages infrequently place only 1.6 percentiles lower, on average.

Figure 6 shows that older immigrants are less likely to feel discriminated based on their origin. The downward life-cycle trajectory of perceived discrimination looks even steeper on the right panel where the age profiles are estimated with individual and year fixed effects. A similar result of greater discrimination against younger immigrants is found in other European OECD countries (OECD, 2013). Thinking along the line of Altonji and Pierret (2001), the trend in Figure 6 could be explained by employer learning under statistical discrimination. As firms learn more about the true productivity of immigrants, the price differential is likely to fall with age. The downward trend could also be explained by the decision of some immigrants to leave the host country after encountering discrimination. Regardless of the reason, the decline in perceived discrimination with age is inconsistent with the price channel of wage divergence. Not only does the trend go counter to the trajectory of the native-immigrant wage gap, which is increasing with age (see Figure 2), it implies that wages should be converging based on the price factor alone. Given the imperfect measure of intertemporal discrimination, we cannot fully refute the price/discrimination channel of wage divergence. Perhaps more targeted data or different research design can shed more light on the role of price discrimination in wage divergence, but the current descriptive evidence negates this channel.

6. Conclusions

Our study finds strong evidence of the wage divergence between natives and immigrants in Germany. Wage divergence is observed despite increasing average wages of immigrants after their arrival in the host country. We introduce two channels of wage divergence into the classic theoretical model of immigrant economic assimilation. The first channel is the efficiency edge of natives over immigrants in the production of host-country-specific human capital. The second channel is differentially changing prices per unit of human capital favoring natives. The theoretical model also predicts that the rate of convergence increases with lower skill transferability, higher valuation of future earnings, more post-migration investment in human capital, and higher (lower) preexisting skills if they are complementary to (substitutable with) post-migration human capital.

Empirical estimates are mostly in line with these predictions. The divergence of relative skill-specific wages does not appear to be attributed to the selection effects arising from survey attrition, out-migration, selection into employment, non-response, and other sources of missing wage growth. After accounting for the endogeneity of post-migration investment in human capital, we find that education in the host country positively contributes to wage convergence. Additional years of schooling in the home country also help immigrants to catch-up with the wages of comparable natives, but the relative contribution of pre-migration investment is considerably smaller than the contribution of post-migration education.

The analysis shows that the proximity of home-country language to host-country language – a common proxy for skill transferability – has a reducing effect on the economic assimilation of immigrants. At the same time, several other proxies that have been traditionally used to measure skill transferability, such as GDP per capita in the home country, ethnicity, and the size of ethnic networks are estimated to have the positive effect on wage convergence. While these last results are not consistent with the skill transferability explanation alone, they make sense according to the efficiency argument. Furthermore, the fact that wage divergence is highest during the first years of work career when people keep investing in human capital also fits well the efficiency story. Several other estimates seem to be supportive of the efficiency hypothesis. For example, we find that ethnic German resettlers and other immigrants who are closer to natives in language, culture, geographic distance, and the level of economic development of their home country receive larger gains from post-migration education in the host country.

The price channel of wage divergence implies the presence of intertemporal discrimination against immigrants, when the price of same skills changes differentially with age by benefiting natives. The available subjective measure of perceived discrimination against immigrants is highly correlated with the lower placement of immigrants in the native wage distribution. However, the age profile of this discrimination variable goes counter to the life-cycle dynamics of the native-immigrant wage gap. If true discrimination follows the same downward trajectory as perceived discrimination, it would imply that wages should be converging with age, which we do not observe. In other words, the evidence from our study does not support the price/discrimination explanation of wage divergence.

Among other notable findings, it is worth mentioning that the secular trend in wage convergence is tilting slightly upward. In contrast to the U.S., where more recent cohorts have smaller rates of economic assimilation (Borjas, 2015), we do not find the slowdown of convergence rates in Germany over time. Furthermore, if the wage growth of immigrants is highly procyclical, the wage convergence is not. Another interesting result is higher assimilation rates among immigrants who flee political violence in the home country compared to those who come from politically stable countries. Also noteworthy is the statistically significant positive effect of the size of well-established ethnic networks on wage convergence. No such effect is found for more recent networks.

Our analysis suggests that greater attention needs to be given to individual differences in assimilation trajectories. Despite the observed wage divergence for an average immigrant, the substantial share of immigrants integrate well into the host-country labor market and catch-up with natives, while many others lag substantially behind. Understanding the individual differences can be crucial in the successful implementation of various integration policies.

References

Adsera, Alicia and Pytlikova, Mariola, 2015. “The Role of Language in Shaping International Migration,” *Economic Journal* 125(586): F49–F81.

Altonji, Joseph G. and David Card, 1991. “The Effects of Immigration on the Labor Market Outcomes of Less-Skilled Natives,” in *Immigration, Trade and the Labor Market*, edited by John M. Abowd and Richard B. Freeman, Chicago: University of Chicago Press: 201-234.

Altonji, Joseph G. and Charles R. Pierret, 2001. “Employer Learning and Statistical Discrimination,” *Quarterly Journal of Economics* 116(1): 313-350.

Angrist, Joshua D, 2001. “Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors,” *Journal of Business & Economic Statistics* 19(1): 2-28.

Basilio, Leilanie, Thomas K. Bauer, and Mathias Sinning, 2009. “Analyzing the Labor Market Activity of Immigrant Families in Germany,” *Labour Economics* 16: 510-520.

Basilio, Leilanie, Thomas K. Bauer, and Anica Kramer, 2017. “Transferability of Human Capital and Immigrant Assimilation: An Analysis for Germany,” *Labour*, forthcoming.

Bellemare, Charles, 2007. “A Life-Cycle Model of Outmigration and Economic Assimilation of Immigrants in Germany,” *European Economic Review* 51: 553-576.

Bollinger, Christopher R. and Barry T. Hirsch, 2006. “Match Bias from Earnings Imputation in the Current Population Survey: The Case of Imperfect Matching,” Journal of Labor Economics 24(3): 483-519.

Borjas, George J., 1995. "Assimilation and Changes in Cohort Quality Revisited: What Happened to Immigrant Earnings in the 1980s?" *Journal of Labor Economics* 13(2): 201-245.

Borjas, George J., 1999. "Economic Analysis of Immigration," in *Handbook of Labor Economics*, edited by Orley Ashenfelter and David Card, Elsevier Science, Vol. 3: 1697-1760.

Borjas, George J., 2015. "The Slowdown in the Economic Assimilation of Immigrants: Aging and Cohort Effects Revisited Again," *Journal of Human Capital* 9(4): 483-517.

Bratsberg, Bernt and James F. Ragan, 2002 "The Impact of Host-Country Schooling on Earnings – A Study of Male Immigrants in the United States," *Journal of Human Resources* 37 (1): 63-105.

Center for Systemic Peace, 2015. *Major Episodes of Political Violence, 1946-2014*, http://www.systemicpeace.org/inscrdata.html

Chin, Aimee and Kalena Cortes, 2015. "The Refugee/Asylum Seeker," in in *Handbook of the Economics of International Migration*, edited by Barry R. Chiswick and Paul W. Miller, North Holland, Vol. 1A: 585-658.

Chiswick, Barry R., 1978. "The Effect of Americanization on the Earnings of Foreign-Born Men," *Journal of Political Economy* 86(5): 897-921.

Chiswick, Barry R. and Paul W. Miller, 1994. "The Determinants of Post-Immigration Investments in Education," *Economics of Education Review* 13(2): 163-177

Chiswick, Barry R. and Paul W. Miller, 2012. "Negative and Positive Assimilation, Skill Transferability, and Linguistic Distance," *Journal of Human Capital* 6(1): 35-55.

Constant, Amelie and Douglas S. Massey, 2003. "Self-selection, Earnings, and Out-migration: A Longitudinal Study of Immigrants to Germany," *Journal of Population Economics* 16: 631-653.

Constant, Amelie F., Simone Schüller, and Klaus F. Zimmermann, 2013. "Ethnic Spatial Dispersion and Immigrant Identity," *IZA Discussion Paper*, No. 7868.

Cortes, Kalena E., 2004. "Are Refugees Different from Economic Immigrants? Some Empirical Evidence on the Heterogeneity of Immigrant Groups in the United States," *Review of Economics and Statistics*. 86(2): 465–480.

Damm, Anna Piil, 2009. "Ethnic Enclaves and Immigrant Labor Market Outcomes: Quasi-Experimental Evidence," *Journal of Labor Economics* 27(2): 281-314.

Duleep, Harriet Orcutt, 2015. "The Adjustment of Immigrants in the Labor Market," in *Handbook of the Economics of International Migration*, edited by Barry R. Chiswick and Paul W. Miller, North Holland, Vol. 1: 105-182.

Duleep, Harriet Orcutt and Regets, Mark C., 1999. "Immigrants and Human-Capital Investment," *American Economic Review* 89(2): 186-191.

Dustmann, Christian and Albrecht Glitz, 2011. "Migration and Education," in *Handbook of the Economics of Education*, edited by Eric A. Hanushek, Stephen Machin and Ludger Woessmann, Elsevier, Vol.4: 327-439.

Dustmann, Christian, Albrecht Glitz, and Thorsten Vogel, 2010. "Employment, Wages, and the Economic Cycle: Differences between Immigrants and Natives," *European Economic Review* 54: 1–17.

Dustmann, Christian and Joseph-Simon Görlach, 2015. "Selective Out-Migration and the Estimation of Immigrants' Earnings Profiles," in *Handbook of the Economics of International Migration*, edited by Barry R. Chiswick and Paul W. Miller, Elsevier, Volume 1A: 489-533.

Edin, Per-Anders, Fredriksson Peter, and Åslund Olof, 2003. "Ethnic Enclaves and the Economic Success of Immigrants: Evidence from a Natural Experiment," *Quarterly Journal of Economics* 118(1): 329-357.

Ethnologue, 2016. *Languages of the World*, 19th edition http://www.ethnologue.com

Ferrer, Ana, David A. Green, and W. Craig Riddell, 2006. "The Effect of Literacy on Immigrant Earnings," *Journal of Human Resources* 41(2): 380-410.

Fertig, Michael and Stefanie Schurer, 2007. "Labour Market Outcomes of Immigrants in Germany: The Importance of Heterogeneity and Attrition Bias," *IZA Discussion Paper*, No. 2915.

Friedberg, Rachel M., 2000. "You Can't Take It with You? Immigrant Assimilation and the Portability of Human Capital," Journal of Labor Economics 18(2): 221-251.

Hatton, Timothy J. and Andrew Leigh, 2011. "Immigrants Assimilate as Communities, Not Just as Individuals," *Journal of Population Economics* 24: 389–419.

Hirano, Keisuke, Guido Imbens, and Geert Ridder, 2003. "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score," *Econometrica* 71, 1161-1189.

Jain, Apoorva and Klara Sabirianova Peter, 2017. "A Joint Hazard-Longitudinal Model of the Timing of Migration, Immigrant Quality, and Labor Market Assimilation," *IZA Discussion Paper*, No.

Kerr, Sari Pekkala and William R. Kerr, 2011. "Economic Impacts of Immigration: A Survey," *NBER Working Paper Series*, No. 16736.

Lubotsky, Darren, 2007. "Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings," *Journal of Political Economy* 115(5): 820–867.

Munshi, Kaivan, 2003. "Networks in the Modern Economy: Mexican Migrants in the U. S. Labor Market," *Quarterly Journal of Economics* 118(2): 549-599.

OECD, 2013. *International Migration Outlook 2013*, OECD Publishing, Paris. http://dx.doi.org/10.1787/migr_outlook-2013-en.

Rosenbaum, Paul R. and Donald B. Rubin, 1983. "The Central Role of The Propensity Score in Observational Studies for Causal Effects," *Biometrika* 70(1): 41-55.

Sanroma, Esteban, Raul Ramos, and Hipolito Simon, 2015. "How Relevant is the Origin of Human Capital for Immigrant Wages? Evidence from Spain," *Journal of Applied Economics* 18(1): 149-172.

Skuterud, Mikal and Mingcui Su, 2012. "The Influence of Measurement Error and Unobserved Heterogeneity in Estimating Immigrant Returns to Foreign and Host-Country Sources of Human Capital," *Empirical Economics* 43: 1109–1141.

Tumen, Semih, 2016. "The Economic Impact of Syrian Refugees on Host Countries: Quasi-Experimental Evidence from Turkey," *American Economic Review* 106 (5): 456-460.

Venturini, Alessandra and Claudia Villosio, 2008. "Labour-Market Assimilation of Foreign Workers in Italy," *Oxford Review of Economic Policy* 24 (3): 517-541.

Wooldridge, Jeffrey M., 2007. "Inverse Probability Weighted M-Estimation for General Missing Data Problems," *Journal of Econometrics* 141(2): 1281-1301.

Wooldridge, Jeffrey M., 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd edition. Cambridge, MA: MIT Press.

Zibrowius, Michael, 2012. "Convergence or Divergence? Immigrant Wage Assimilation Patterns in Germany," SOEP Papers, # 479.

Table 1: Sample Composition

Panel A: Native-Immigrant Comparison

	Natives	***Immigrants***
Female	0.514	0.532
Age	41.789	45.302
	(13.631)	(10.977)
Adjusted years of schooling	11.169	10.182
	(2.969)	(2.512)
Parents' education		
Basic sec, lower vocational or less	0.458	0.639
General sec and upper vocational	0.383	0.176
Higher education	0.120	0.083
Unknown	0.039	0.101
Urban residence	0.752	0.859
N of observations	247,967	45,769

Panel B: Immigrants Only

	At arrival	***Post-migration at survey time***
Ethnic German	...	0.133
Age at migration	26.257	...
	(7.762)	
Number of years since migration	0.000	19.045
		(10.199)
Any education (schooling + training)	1.000	0.223
Years of education	10.481	0.622
	(2.839)	(1.692)
Any formal schooling	1.000	0.137
Years of formal schooling	9.657	0.354
	(2.649)	(1.302)
Any job training	0.308	0.136
Years of job training	0.824	0.268
	(1.438)	(0.930)

Notes: Standard deviations for non-binary variables are provided in parentheses. Number of immigrants =7,496 [age 17-65, age at migration 15 or older, year of migration 1961-2014, reside in West Germany]; number of natives =31,215 [age 17-65, reside in West Germany]. Other sample constraints are discussed in Appendix A1. Years of schooling in two panels are not comparable. Panel A uses the typical length of study in each completed level of education, while Panel B is based on actual years of schooling reported in the retrospective calendar; see Appendix A2 for details. Summary statistics is reported using sample weights (see Appendix A1).

Table 2: Labor Market Outcomes of Natives and Immigrants

	Log Hourly Wage	***Employment Participation***	***Unemployment Probability***
Natives, mean(se)	2.196	0.706	0.050
	[0.001]	[0.001]	[0.001]
Immigrants, difference			
Unconditional	-0.065	-0.085	0.081
	(0.000)	(0.000)	(0.000)
Conditional on a quartic in age	-0.173	-0.123	0.083
	(0.000)	(0.000)	(0.000)
Conditional on X_{it}	-0.110	-0.100	0.072
	(0.000)	(0.000)	(0.000)
N of observations	159,829	247,967	184,717

Notes: Standard errors for the mean estimate are in square brackets. The p-values for the mean difference t-test are reported in parentheses. The difference is the for the labor market outcomes between natives and immigrants, with positive sign favoring immigrants. In addition to a quartic polynomial in age, the vector X_{it} includes a dummy for being female, adjusted years of schooling, urban residence, and fixed effects for survey year. The unemployment rate is conditional on being in the labor force. The construction of each variable is described in Appendix A2. The estimates use sample weights.

Table 3: Summary Statistics

Binary Variables	*Sample A*	*Sample B*	*Continuous Variables*	*Sample A*	*Sample B*
Post-migration education, [*a*, *t*]	0.231	0.205	Pre-migration years of	10.400	10.236
Post-migration schooling [*a*, *t*]	0.133	0.124	education, *a*	(2.707)	(2.702)
Post-migration job training [*a*, *t*]	0.147	0.127	Pre-migration years of	9.479	9.412
Female	0.425	0.524	formal schooling, *a*	(2.534)	(2.468)
Ethnic German	0.168	0.140	Pre-migration years of	0.920	0.825
Age group, *t*			job training, *a*	(1.475)	(1.430)
26-35	0.178	0.175	Log of GDP per capita	9.199	9.137
36-45	0.335	0.271	in home country, *a*	(0.566)	(0.595)
46-55	0.377	0.302	Ethnic networks, *t*	1.229	1.230
56-65	0.086	0.212		(1.038)	(1.043)
Parents' education			Linguistic proximity	0.223	0.196
Secondary education	0.117	0.126		(0.193)	(0.193)
Higher education	0.063	0.070	GDP per capita growth in	1.517	1.475
Unknown	0.105	0.117	Germany, [*t*+1, *t*+5]	(0.648)	(0.621)
Urban residence, *t*	0.868	0.870	GDP per capita growth in	1.889	2.273
Instability in home country, *a*			home country, [*t*+1, *t*+5]	(3.791)	(3.206)
Limited political violence	0.085	0.091	Average commuting distance, *t*	12.854	13.211
Serious political violence	0.048	0.077		(2.189)	(2.147)
Warfare	0.081	0.115	Political instability score	0.526	0.744
Mode of interview, *t*			in home country, ages 6-10	(1.513)	(1.720)
Self-written and mailed	0.043	0.063	Wage growth, [*t*+1, *t*+5]	0.930	1.040
Computer assisted	0.140	0.163		(7.231)	(17.181)
Interviewer, *t*			Change in conditional relative	-0.069	-0.014
First interview	0.084	0.094	relative wage, [*t*+1, *t*+5]	(5.509)	(10.937)
Different interviewer	0.090	0.090	*N of observations*	13,353	37,253

Notes: Table shows the mean and standard deviation of variables in the two samples of immigrants: (A) censored sample with non-missing wage growth data and (B) combined censored and uncensored sample, which includes immigrants with missing wage growth data. Both samples end in 2009, which is the last year used to compute wage growth in the next 5-year period (the main reason for the difference with Table 1). Standard deviations are in parenthesis and not reported for dummy variables. Superscript *a* indicates the year of arrival, and *t* denotes the year of survey. Base/omitted categories are age 17-25 for age groups, "basic secondary and lower vocational" for parents' education, "no episodes of political violence" for instability in home country, "face-to-face interviews" for the mode of interview and the same interviewer. All variables are described in Appendix A2. Summary statistics is reported using sample weights (see Appendix A1).

Table 4: Wage Convergence Equation

	OLS	*IPW*	*IMR*	*MLE*
Pre-migration years of education	0.053**	0.016	0.055**	0.054**
	(0.021)	(0.030)	(0.023)	(0.021)
Female	0.257***	0.352***	0.174	0.233**
	(0.096)	(0.124)	(0.207)	(0.109)
Ethnic German	0.317**	0.287	0.346**	0.325**
	(0.139)	(0.176)	(0.156)	(0.140)
Age group				
26-35	1.525***	1.945***	1.587***	1.543***
	(0.382)	(0.446)	(0.384)	(0.383)
36-45	1.190***	1.602***	1.269***	1.213***
	(0.378)	(0.440)	(0.395)	(0.381)
46-55	1.175***	1.619***	1.240***	1.194***
	(0.384)	(0.446)	(0.390)	(0.386)
55-65	2.431***	2.853***	2.304***	2.395***
	(0.429)	(0.497)	(0.552)	(0.439)
Parents' education				
General sec and upper vocational	-0.142	0.403*	-0.166	-0.149
	(0.155)	(0.214)	(0.165)	(0.155)
Higher education	0.960***	1.310***	0.939***	0.954***
	(0.279)	(0.453)	(0.293)	(0.280)
Urban residence	-0.313**	-0.597***	-0.308**	-0.311**
	(0.138)	(0.184)	(0.134)	(0.137)
Log of GDP per capita in	0.340**	0.338*	0.341**	0.340**
home country at arrival	(0.146)	(0.195)	(0.151)	(0.146)
Instability in home country at arrival				
Limited political violence	-0.392**	-0.063	-0.407**	-0.396**
	(0.161)	(0.232)	(0.171)	(0.162)
Serious political violence	0.211	-0.223	0.146	0.192
	(0.345)	(0.417)	(0.381)	(0.349)
Warfare	0.597***	0.537**	0.571**	0.590***
	(0.218)	(0.240)	(0.222)	(0.219)
Ethnic networks	0.202***	0.225**	0.203***	0.202***
	(0.071)	(0.101)	(0.071)	(0.071)
Linguistic proximity	-0.997***	-1.070**	-0.872*	-0.961***
	(0.360)	(0.478)	(0.452)	(0.370)
Intercept	-5.109***	-5.044***	-5.443***	-5.205***
	(1.464)	(1.928)	(1.712)	(1.484)
Selection term	...	...	0.278	0.015
			(0.656)	(0.037)
Year FE	Yes	Yes	Yes	Yes
Wald χ_2			348.5***	279.6***

Notes: N=13,353. Table presents estimates of the reduced-form wage convergence model for immigrants using different treatments of selectivity bias: inverse propensity weighting (IPW), the Heckman 2-step procedure with the inverse Mills ratio (IMR), and the full maximum likelihood estimation (MLE). Standard errors are in parentheses; robust for OLS, IPW, and MLE; bootstrapped with 1000 iterations for IMR; *** p<0.01, ** p<0.05, * p<0.1. Base/omitted categories are shown in notes to Table 3. Unknown parents' education is also included in the estimates but not shown here. The dependent variable is the annual change in immigrants' conditional relative wage averaged over the 5-year period. The selection term is the coefficient on the inverse Mills ratio in IMR and the estimated rho in MLE.

Table 5: Selection Probit Equation, Marginal Effects

Variables	***ME***	***Variables***	***ME***
		Instability in home country at	
Pre-migration years of education	0.004***	arrival	
	(0.001)	Limited political violence	-0.029***
Female	-0.149***		(0.008)
	(0.005)	Serious political violence	-0.109***
Ethnic German	0.064***		(0.012)
	(0.007)	Warfare	-0.050***
Age group			(0.009)
26-35	0.093***	Ethnic networks	-0.005
	(0.012)		(0.003)
36-45	0.121***	Average commuting distance	-0.007***
	(0.012)		(0.002)
46-55	0.096***	GDP per capita growth in home	-0.005***
	(0.012)	country, 5-year MA [*t*+1, *t*+5]	(0.001)
55-65	-0.219***	Mode of interview	
	(0.013)	Self-written and mailed	-0.112***
Parents' education			(0.010)
General sec and upper vocational	-0.030***	Computer assisted	-0.000
	(0.008)		(0.009)
Higher education	-0.033***	Interviewer	
	(0.012)	First interview	-0.083***
Urban residence	0.010		(0.011)
	(0.007)	Different interviewer	-0.020**
Log of GDP per capita in	0.001		(0.008)
home country at arrival	(0.006)	Intercept	Yes
Linguistic proximity	0.198***	Year FE	Yes
	(0.019)	*N of observations*	37,253

Notes: Table presents estimates of the selection probit equation described in Section 4.3. Reported are the marginal effects (MEs) evaluated at sample means. The means can be found in column "Sample B" of Table 3. Robust standard errors are in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Base/omitted categories are shown in notes to Table 3. Unknown parents' education is also included in the estimates but not shown here. The estimates of the selection equation from the MLE specification are similar and not reported.

Table 6: Wage Growth Equation

	OLS	*IPW*	*IMR*	*MLE*
Pre-migration years of education	0.017	0.060	0.022	0.018
	(0.029)	(0.043)	(0.032)	(0.029)
Female	0.501***	0.568***	0.287	0.456***
	(0.131)	(0.168)	(0.291)	(0.140)
Ethnic German	0.464**	0.352	0.539**	0.480**
	(0.201)	(0.285)	(0.233)	(0.203)
Age group				
26-35	-2.626***	-2.311***	-2.469***	-2.593***
	(0.508)	(0.602)	(0.518)	(0.509)
36-45	-4.124***	-3.867***	-3.921***	-4.081***
	(0.502)	(0.592)	(0.546)	(0.505)
46-55	-3.977***	-3.737***	-3.809***	-3.942***
	(0.507)	(0.594)	(0.530)	(0.509)
55-65	-4.179***	-4.016***	-4.506***	-4.247***
	(0.553)	(0.664)	(0.716)	(0.558)
Parents' education				
General sec and upper vocational	0.132	0.701**	0.072	0.119
	(0.224)	(0.314)	(0.235)	(0.224)
Higher education	1.760***	2.401***	1.707***	1.749***
	(0.405)	(0.677)	(0.432)	(0.405)
Urban residence	-0.567***	-0.871***	-0.555***	-0.565***
	(0.174)	(0.223)	(0.173)	(0.173)
Log of GDP per capita in	0.667***	0.710***	0.670***	0.668***
home country at arrival	(0.205)	(0.275)	(0.208)	(0.204)
Instability in home country at arrival				
Limited political violence	-0.399*	-0.049	-0.438*	-0.407*
	(0.216)	(0.308)	(0.224)	(0.216)
Serious political violence	0.568	0.167	0.401	0.533
	(0.455)	(0.572)	(0.503)	(0.456)
Warfare	0.620*	0.627	0.553*	0.606*
	(0.324)	(0.386)	(0.334)	(0.324)
Ethnic networks	0.218**	0.341**	0.221**	0.219**
	(0.092)	(0.137)	(0.092)	(0.092)
Linguistic proximity	-1.613***	-1.730**	-1.290**	-1.545***
	(0.495)	(0.697)	(0.644)	(0.504)
Intercept	-0.293	-1.472	-1.148	-0.472
	(2.058)	(2.746)	(2.412)	(2.071)
Selection term	...	...	0.713	0.021
			(0.942)	(0.029)
Year FE	Yes	Yes	Yes	Yes
Wald χ_2			507.4***	498.2***

Notes: N=13,353. Table presents estimates of wage growth equation using different treatments of selectivity bias: inverse propensity weighting (IPW), the Heckman 2-step procedure with the inverse Mills ratio (IMR), and the full maximum likelihood estimation (MLE). Standard errors are in parentheses; robust for OLS, IPW, and MLE; bootstrapped with 1000 iterations for IMR; *** p<0.01, ** p<0.05, * p<0.1. Base/omitted categories are shown in notes to Table 3. Unknown parents' education is also included in the estimates but not shown here. The dependent variable is annual wage growth averaged over the 5-year period. The selection term is the coefficient on the inverse Mills ratio in IMR and the estimated rho in MLE.

Table 7: Alternative Specifications

	Wage Convergence		*Wage Growth*	
	OLS	***MLE***	***OLS***	***MLE***
Panel A				
GDP per capita growth in Germany	-0.115	-0.122	0.425***	0.421***
5-year MA [t+1, t+5]	(0.080)	(0.080)	(0.108)	(0.108)
Linear time trend	0.018*	0.012	-0.062***	-0.065***
	(0.009)	(0.010)	(0.013)	(0.013)
Panel B				
Pre-migration years of formal schooling	0.079***	0.080***	0.045	0.046
	(0.025)	(0.025)	(0.035)	(0.035)
Pre-migration years of job training	-0.003	-0.003	-0.041	-0.041
	(0.033)	(0.033)	(0.041)	(0.041)
Year FE	Yes	Yes	Yes	Yes
Panel C				
Ethnic networks - established	0.209***	0.209***	0.264***	0.264***
	(0.075)	(0.075)	(0.097)	(0.096)
Ethnic networks - recent	0.125	0.132	-0.248	-0.234
	(0.199)	(0.199)	(0.259)	(0.259)
Year FE	Yes	Yes	Yes	Yes

Notes: N=13,353. Table shows additional specifications of the wage convergence and wage growth models presented in Tables 4 and 6. In Panel A, year dummies are replaced with a linear trend and country-level economic growth in Germany. In Panel B, pre-migration years of education are divided between formal schooling and job training. In Panel C, ethnic networks are split into established and recent networks. In every other way, these three specifications are identical to the ones reported in Tables 4 and 6. Robust standard errors are in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The dependent variables are (i) annual change in the immigrant position in the native conditional wage distribution and (ii) annual wage growth, both are averaged over the 5-year period.

Table 8: Treatment Effect of Post-Migration Education on Wage Convergence

	N	*ATE*	*ATT*
	Panel A: Overall		
IPWRA1	13,144	0.540***	0.468***
		(0.203)	(0.144)
IPWRA2	13,259	0.368*	0.365**
		(0.191)	(0.143)
Simple IPW	13,146	0.746***	0.355**
		(0.278)	(0.141)
Augmented IPW	13,146	0.638**	...
		(0.280)	
	Panel B: IPWRA1 By Group		
Ethnic Germans	3,121	0.492**	0.396*
		(0.213)	(0.213)
Non-ethnic Germans	9,549	0.051	-0.006
		(0.280)	(0.216)
Same family language	9,600	1.006***	0.442**
		(0.217)	(0.176)
Different family language	3,407	-0.237	0.145
		(0.389)	(0.259)
More developed country	6,608	0.663***	0.463**
		(0.234)	(0.192)
Less developed country	6,316	0.394	0.330
		(0.366)	(0.222)
Neighboring country	1,743	0.805***	0.783**
		(0.303)	(0.375)
Non-neighboring country	11,303	0.477**	0.355**
		(0.231)	(0.158)

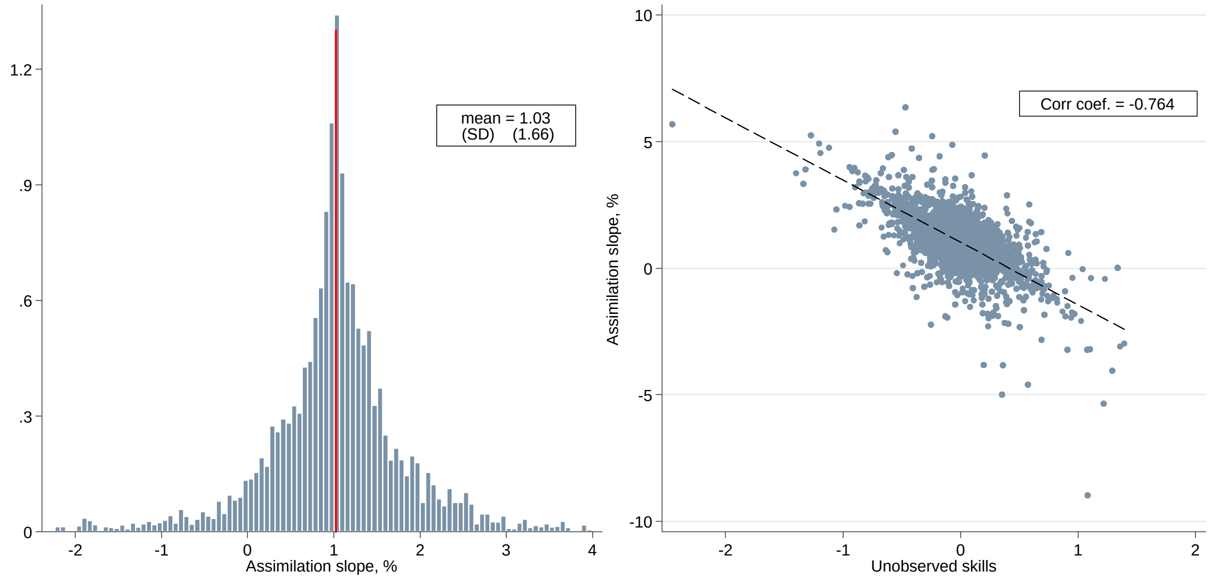
Notes: Table presents the estimated treatment effects of post-migration education on immigrants' wage convergence. ATE=average treatment effect; ATT=average treatment effect on the treated; IPWRA=doubly robust inverse-probability weighted regression adjustment (Wooldridge, 2010). The method is described in Section 4.4. IPWRA1 includes IVs and a quadratic trend. IPWRA2 excludes IVs but includes year fixed effects. The set of auxiliary equations used in calculating the treatments effects in line 1 is reported in Appendix Table W3. The sum of *N* varies because observations with the propensity score outside the range [0.001, 0.999] in each group are dropped to satisfy the overlap conditions. Robust standard errors are in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Subsamples are described in Appendix Table A2.

Table 9: IV Effect of Post-Migration Education on Wage Convergence

	N	*OLS*	*Probit-2SLS*	*Probit-GMM*	*Probit-LIML*
Overall	13,353	0.276**	3.241**	3.044*	3.750**
		(0.129)	(1.561)	(1.554)	(1.848)
Ethnic Germans	3,126	0.288	5.248*	5.346**	7.699*
		(0.185)	(2.722)	(2.718)	(4.393)
Non-ethnic Germans	10,227	0.103	4.585**	4.139**	4.956**
		(0.178)	(1.867)	(1.846)	(2.033)
Same family language	9,887	0.437***	8.143***	7.695***	8.865***
		(0.148)	(2.719)	(2.661)	(3.058)
Different family language	3,466	-0.086	-2.277	-2.176	-2.361
		(0.270)	(2.352)	(2.341)	(2.444)
More developed country	6,653	0.411**	11.287***	11.219***	11.579***
		(0.167)	(4.094)	(4.067)	(4.255)
Less developed country	6,700	0.100	0.695	0.436	0.723
		(0.206)	(1.187)	(1.183)	(1.241)
Neighboring country	1,748	0.560**	3.021	3.883	6.677
		(0.272)	(2.430)	(2.439)	(6.408)
Non-neighboring country	11,605	0.199	0.747	0.398	0.831
		(0.146)	(1.281)	(1.276)	(1.475)

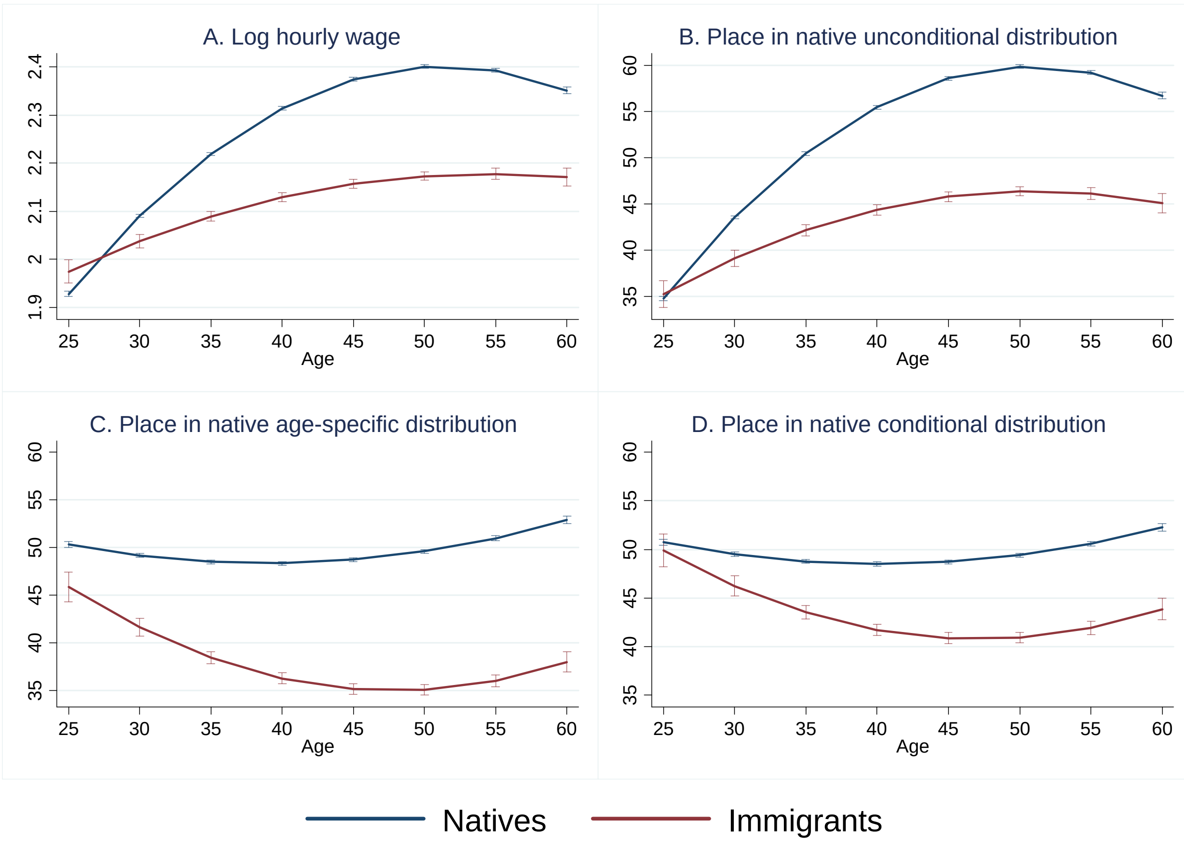
Notes: Table presents the estimated effect of post-migration education on immigrants' wage convergence using OLS and IV approaches. The method is described in Section 4.4. IVs include the log of the supply of training offers, the excess demand for training, and political instability in home country at ages 6-10. 2SLS=two-stage least squares; GMM=generalized method of moments; LIML=limited-information maximum likelihood. Full results from the OLS and IV estimation in line 1, including the first stage probit, are reported in Appendix Table W4. Robust standard errors are in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Subsamples are described in Appendix Table A2.

Figure 1: Individual Returns to Years Since Migration



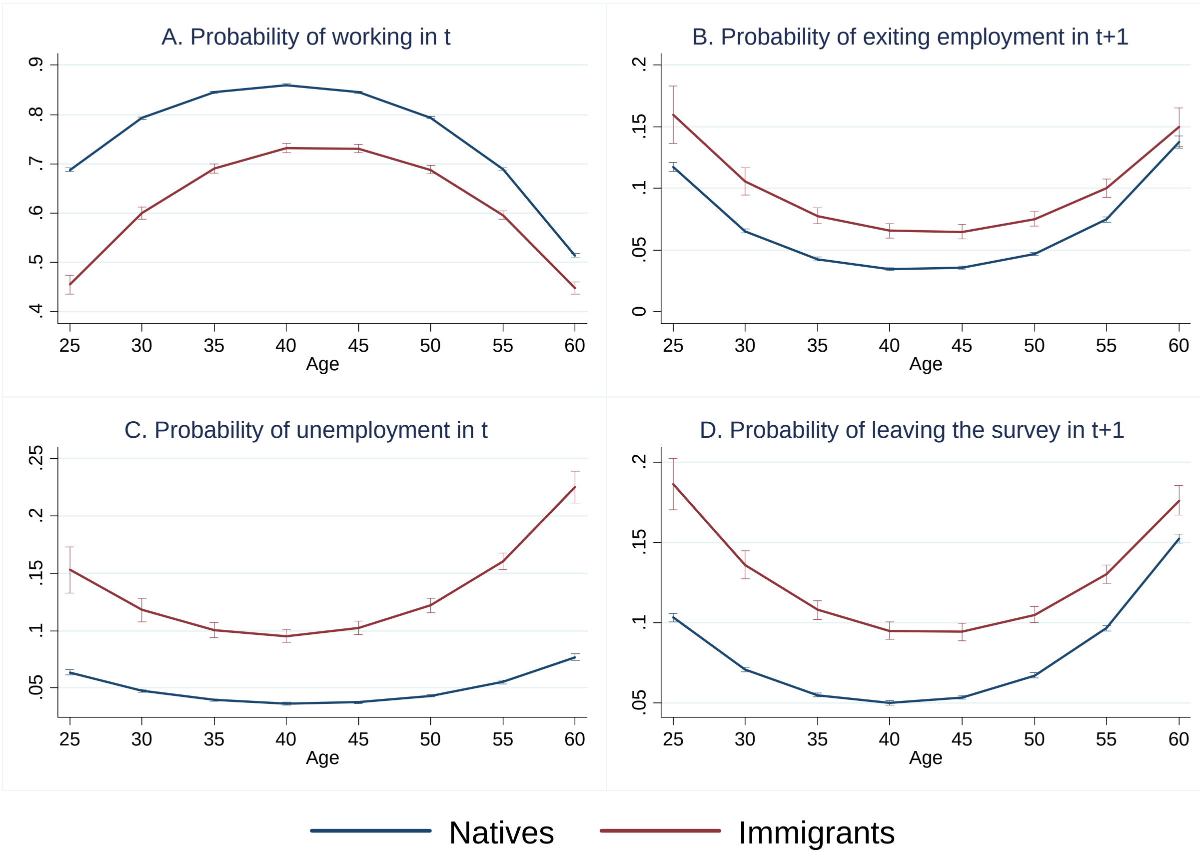
Notes: The left panel shows the distribution of returns to years since migration. The returns are obtained from the linear mixed model with correlated random intercepts and slopes, as shown in Equation (1). The dependent variable is the log of hourly wage. The plotted returns are the best linear unbiased predictors (BLUPs) of the random slope of the number of years since migration plus the mean return. When multiplied by 100, they show the percent wage increase for each additional year spent in the host country. The panel on the right shows the scatterplot of the BLUPs of random slope and random intercept estimated from the same linear mixed model.

Figure 2: Age-Wage Profiles: Evidence of Wage Divergence



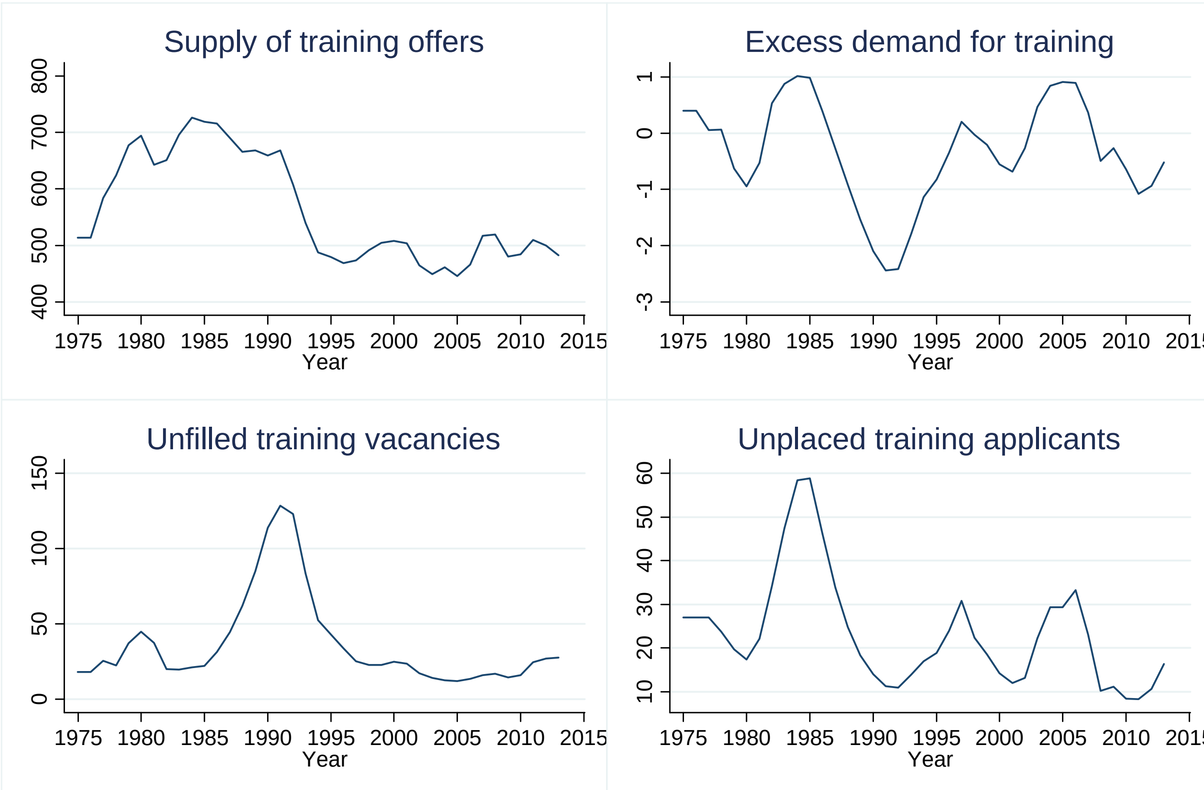
Notes: The life-cycle profiles of the log of hourly wage and relative wages are predicted marginal effects from the OLS regression of the corresponding outcome for a given group (natives or immigrants) on a quadratic polynomial in age with robust standard errors. The relative wage is defined as the placement of immigrants in the native wage distribution. Definitions of relative wage are discussed in Section 2. The 95 percent confidence interval for the point estimate is also shown.

Figure 3: Employment Outcomes and Survey Participation over Life Cycle



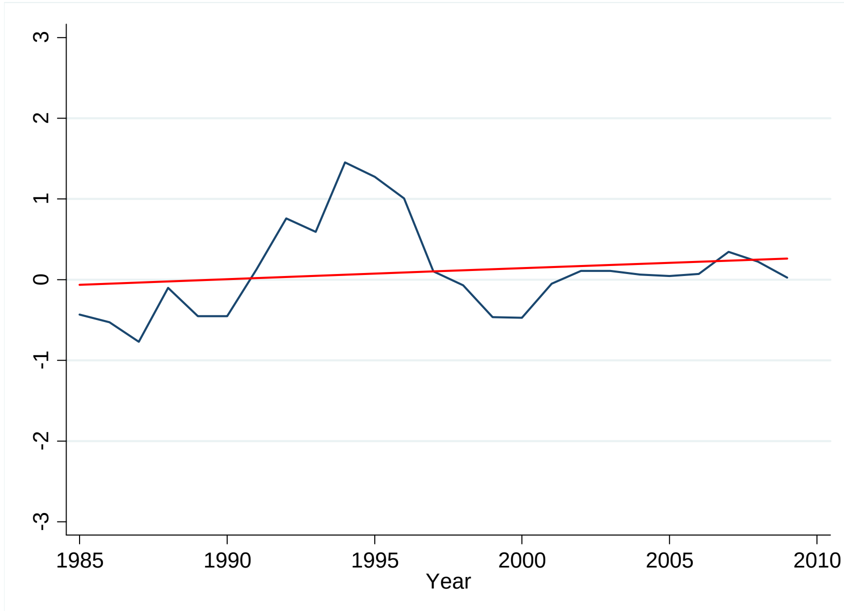
Notes: The trajectories are predicted marginal effects from the probit regression of the corresponding outcome for a given group (natives or immigrants) on a quadratic polynomial in age with robust standard errors. The probability of unemployment is conditional on being in the labor force. The probability of exiting employment in $t+1$ is conditional in being employed in t. The migration status is a binary variable that takes the value of 0 for natives and 1 for immigrants. The 95 percent confidence interval for the point estimate is also shown.

Figure 4: Demand and Supply of Training Programs



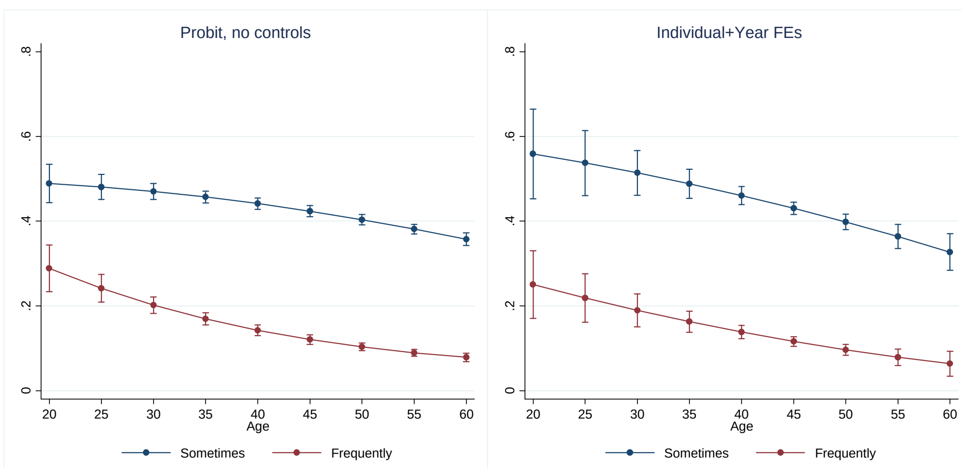
Notes: This figure plots the time-series for the demand and supply of government-sponsored training programs in West Germany. The excess demand is calculated as the difference between the log of unplaced training applicants and the log of unfilled training places still registered with employment offices. Other indicators are in thousands of people/vacancies per year. The data source is the German Federal Ministry of Education and Research.

Figure 5: Year Effects and Secular Trend in Wage Convergence



Notes: This figure plots the year effects from the estimated model of wage convergence presented in Table 4, Column 4 (MLE). The red line shows a secular trend in wage convergence.

Figure 6: Perceived Discrimination by Age



Notes: The trajectories are predicted marginal effects from the probit regression of the corresponding outcome on a quadratic polynomial in age. The estimates on the right panel also include individual and year fixed effects. The outcomes "sometimes" and "frequently" are the answers of immigrants to the question "How often have you personally felt disadvantaged in Germany in the last two years because of your origins?" The zero value is assigned if the answer is "never". See Appendix A2 on how two binary variables are constructed. The 95 percent confidence interval for the point estimate with robust standard errors is also shown. The estimates are for the 1996-2011 and 2013 surveys. Number of observations is 18,378.

Web appendix for paper "Limits to Wage Growth: Understanding the Wage Divergence between Immigrants and Natives"

Appendix A1. Data

The German Socio-Economic Panel (GSOEP) is the longest-running panel of private households and persons in Europe. It is widely used in the migration research, as it is one of a few national longitudinal surveys with a large representation of immigrants and substantial information on immigrants. Some examples of the published migration studies based on GSOEP include Brücker *et al.* (2014b), Constant *et al.* (2009), Jaeger *et al.* (2010), Zimmermann (2007), among others.

A. Samples

GSOEP is collected and distributed by the German Institute for Economic Research, DIW Berlin. The survey started in 1984 and includes 31 survey waves as of 2014. In 1990, residents of the German Democratic Republic (East Germany) were included in the target population. Later, several additional samples were drawn to replenish the original sample and to include special sub-populations such as immigrants and high-income households.

With respect to immigrants, all samples can be divided into two large groups. In the first group (samples, A, C, E, G, H, J, and K), immigrants are sampled as part of the total population or subpopulation. In the random samples of total population, the share of immigrants is relatively low: about 4 percent in the initial sample A and 7 to 12 percent in replenishment samples E, H, J, and K. Immigrants constitute about 5 percent of high-income earners in sample G and a mere 1.3 percent of East Germans in sample C. The second group of GSOEP samples includes samples B, D, and M, which focused specifically on immigrants. Sample B "Foreigners in West Germany" started in 1984 with 1393 households whose head came from one of the five largest guest-worker countries (Turkey, Italy, the former Yugoslavia, Greece, and Spain). Sample D "Immigrants" started in 1994/95 with 522 households, which consisted primarily of ethnic German immigrants from the former Soviet Union and Eastern Europe as well as asylum seekers mainly from the parts of Yugoslavia devastated by the war. Finally, sample M "Migration" started in 2013 with 2,723 households. It is designed to account for changes in the composition of migration to Germany since 1995 (Brücker *et al.*, 2014b). All three migration-focused samples have a substantial share of native-born population (27 percent), since some members of households are born in Germany.

The immigrant status in GSOEP is defined based on the country of birth or in earlier waves based on the country of residence since 1949. Immigrants in GSOEP come from more than 130 countries. We are using the 95 percent of the original data since researchers outside the European Union are not allowed to access the entire dataset. We limit the sample of immigrants to those who were between the ages of 17 and 65 at the time of survey and who arrived to Germany after 1960 at age 14 or older. The sample of natives is constrained by age 17-65. We drop observations with missing values on migration status, country of origin, the year of migration, work experience, and the level of schooling. In total, we drop 4.6 percent of observations with missing values in the constrained sample. Given a very small percent of missing values, we assume that dropped observations are ignorable or missing completely at random.

B. Weights

Immigrant-focused GSOEP samples are not a random draw from the German immigrant population, and their composition in the GSOEP does not match the national composition of immigrants by country of origin. Due to sampling design, there is a substantial oversampling of immigrants from the countries that signed guest-worker agreements and also from Poland and former Soviet Union. At the same time, immigrants from Asia, Africa, Middle East and other geographic areas are under-sampled. It is apparent that sample re-weighting is required to match the sample moments to the population moments. The GSOEP provides researchers with cross-sectional weights **phrf*, which we renamed as ***CWEIGHT***. We employ these weights in calculating national or regional averages such as moments of national wage distribution or average commuting distance from home to workplace. However, since many immigrants are sampled outside the main sampling frame, their cross-sectional weight is often set to zero; for example, more than 40 percent of sample D “Immigrants” have zero sampling weight.

To keep as many surveyed immigrants as possible in our estimation sample, we develop immigrant sampling weights (***IWEIGHT***) based on the annual share of each home country in total German population. The OECD International Migration Database (OECD, 2016a) and the German Central Register of Foreign Nationals (Ausländerzentralregisters) report the annual composition of foreign population by origin, which covers more than 99 percent of foreign population from 1990 to 2015 and 91 to 96 percent from 1984 to 1989. *IWEIGHT* is obtained as a ratio of the country share in total German population to the country share in GSOEP sample for each year separately. The *IWEIGHT* for German-born respondents is above 1 due to oversampling of the immigrant population in GSOEP; it ranges between 1.01 and 1.24, with mean=1.08.

Most oversampled home countries with *IWEIGHT* below 0.3 are countries of the former Soviet Union and Poland, while the top under-sampled countries with weights above 2 are Israel, Australia, and countries of East Asia and the Pacific. In addition to using *IWEIGHT*, we control for standard weighting factors such as gender, age, and urban area in model estimates.

C. References for GSOEP data description

Ausländerzentralregisters (Central Register of Foreign Nationals). Ausländische Bevölkerung (Foreign Population Statistics), Statistisches Bundesamt (Federal Statistical Office), annual.

Brücker, Herbert, Andreas Hauptmann, Elke J. Jahn, and Richard Upward, 2014a. "Migration and Imperfect Labor Markets: Theory and Cross-country Evidence from Denmark, Germany and the UK," *European Economic Review*, Volume 66, February 2014: 205-225.

Brücker, Herbert, Martin Kroh, Simone Bartsch, Jan Goebel, Simon Kühne, Elisabeth Liebau, Parvati Trübswetter, Ingrid Tucci, and Jürgen Schupp, 2014b. "The New IAB-SOEP Migration Sample: An Introduction into the Methodology and the Contents," *SOEP Survey Papers*, No. 216: Series C. Berlin: DIW/SOEP.

Constant, Amelie F., Liliya Gataullina, and Klaus F. Zimmermann, 2009. "Ethnosizing Immigrants," *Journal of Economic Behavior & Organization*, 69(3), March 2009: 274-287.

Jaeger, David A., Thomas Dohmen, Armin Falk, David Huffman, Uwe Sunde, and Holger Bonin, 2010. "Direct Evidence on Risk Attitudes and Migration," *Review of Economics and Statistics*, August 2010, 92(3): 684-689.

Kroh, Martin, Simon Kühne, Rainer Siegers, 2015. "Documentation of Sample Sizes and Panel Attrition in the German Socio Economic Panel (SOEP) (1984 until 2014)," *SOEP Survey Papers*, No. 297: Series C. Berlin: DIW / SOEP.

OECD, 2016a. OECD International Migration Database, http://www.oecd.org/els/mig/keystat.htm

Zimmermann, Klaus F., 2007. "The Economics of Migrant Ethnicity," *Journal of Population Economics* 20: 487-494.

Appendix A2. Variables

A. Individual-Level Variables

Country of origin and year of immigration

Country of origin is defined as Germany if a person is born in Germany or immigrated before 1949. Other 130+ countries of origin are re-coded according to the UN country classification in order to link individual observations with macro indicators. Kurdistan is coded as Turkey, Benelux as Netherlands, and the Free City of Gdansk as Poland. Categories for “No nationality”, “Africa”, “Other unspecified foreign country”, and “Unspecified country within EU” are coded as missing. The category “unspecified Eastern Europe”, which mostly includes immigrants from former German territories of Eastern Europe, is kept separately, but linked with macro indicators from Poland. Year of immigration is the calendar year in which the first immigration to territories of the Federal Republic of Germany occurred. Both of these variables are provided for public use as part of the biography and life history data; see documentation of biography variables in SOEP (2014a).

Years since migration (YSM)

Number of years since immigration, or the length of stay in the host country, is calculated as year of survey minus year of immigration.

Female, age, year of survey

Self-explanatory.

Adjusted years of schooling

Adjusted years of schooling reflect the highest level of schooling achieved rather than the total number of years attended in school. Adjusted years are created based on the type of completed secondary school, vocational training, and university education. The following years are assigned for each completed level of secondary schooling: basic secondary school = 9 years; intermediate secondary school = 10 years; technical secondary school = 12 years; academic secondary school = 13 years; other secondary degree = 10 years; dropout = 7 years; currently in secondary school = 7 years.

For those who completed university education, the following years are added to 13 years of academic secondary school: 3 years for technical college, 5 years for university, college abroad, or engineering institute in East Germany, and 8 years for the master or doctorate degree. Finally, 2 more years are added to years of basic or intermediate secondary schooling if the respondent completed 2-year vocational or technical school.

Years of formal schooling and job training

These variables are constructed using the spell dataset on activity status between the ages of 15 and 65; see description in SOEP (2014a). We start with six main activities that include formal schooling, job training, full-time employment, part-time employment, military/civil service, and unemployment. If more than one activity is reported in a given year, then each activity gets a corresponding share of one year. We assume continuous schooling from age 7 to age 14 and no job training before age 15. Then, years of formal schooling and job training at each age are calculated as a running sum of corresponding spells up to a given age.

Based on the age at migration, we construct the number of years of education (schooling and training separately and combined) at the time of arrival to Germany and the number of years studied in Germany after migration up to a given age. Given that the distribution of years spent on investment in post-migration human capital is highly skewed, we dichotomize years into a binary variable for any education received in Germany.

Ethnic German

A dummy variable indicating if an immigrant is of German descent from Eastern Europe. Ethnicity is only available for immigrants.

Parents' education

The variable represents the highest level of schooling completed by a parent: [1] Level I "Basic secondary, lower vocational or less", [2] Level II "General secondary or upper vocational", [3] Level III "Higher education or more", and [4] "Unknown level of parents' education". The first category is chosen as a base category. This variable is constructed based the level of general schooling and the level of professional education provided for each parent in the biography dataset *BIOPAREN* (SOEP, 2014a). First, we aggregate all levels of schooling into three categories. Level III includes degrees from technical engineering school, college, university, and foreign college. Level II includes degrees from intermediate school, technical school, upper secondary school, vocational school, foreign vocational school, health care school, and special technical school. Level I consists of other types of schooling, which are not in Level II or III and include basic secondary school degree, incomplete secondary school, no schooling, apprenticeship, and on-the-job training. Then, we choose the highest level completed among parents. If information is only available for one parent, only that parent's data is used. If the level of schooling is missing for both parents, then these respondents are combined into the fourth category "Unknown level of parents' education". The share of immigrants in the unknown category is about 10 percent.

Urban area

A dummy variable indicating if the respondent resides in urban area. It is part of the HBRUTTO data file in GSOEP. The variable is nearly time-constant; only 2 percent of all immigrants move from urban to rural or back (0.25 percent per annum).

Hourly wage

This variable is based on the net income earned from employment last month in constant 2010 prices (in Euro). Net income means the amount after deduction of taxes, social security, and unemployment and health insurance. The amount excludes vacation pay or back pay. Net labor earnings last month in current prices are part of the dataset of generated variables (SOEP, 2014b). The calculation of the log of hourly wage involves the following steps:

- Exclude imputed values of net labor earnings due to potential match bias from earnings imputation (Bollinger and Hirsch, 2006). Instead, we use the selection-correction methods to account for missing values in earnings.
- Deflate labor earnings to 2010 Euros by using annual CPI for Germany (West Germany until 1990) (OECD, 2016).
- Divide real labor earnings by the product of actual working hours per week and (30/7) number of weeks in a month. Contractual hours are not used because they are not available for the self-employed and exclude over-time work.
- Take the log of the calculated hourly wage and denote it as w_{it}.

Based on the log of hourly wage, w_{it}, we construct the following variables:

- **Wage growth, $\dot{w}_{it}$,** is the average annual growth over the future 5-year period, $\dot{w}_{it} = \left(\frac{1}{5}\right)\sum_{t}^{t+4}(w_{it+1} - w_{it})$. The minimum of three out of five possible growth data points is used in calculating average annual wage growth over the 5-year period. The average annual growth is chosen over the 5-year log difference $(w_{t+5} - w_t)$ to retain information from interim years, to mitigate noise in reported income and hours, and to reduce the influence of temporary jumps/drops in wage rate. The 5-year interval is chosen because it is not too short to be overly sensitive to transitory earnings shocks and to the measurement error, but not too long to lose a significant number of observations due to survey attrition and outmigration.
- **Wage growth observed** is a binary indicator that takes up the value of one if wage growth is observed, and the value of zero if otherwise.
- **Relative wage (percentile), θ_{it},** is the position of the immigrant in the wage distribution of comparable natives with the same observed characteristics. In constructing relative wage, we first obtain the percentile values of the residuals from the regression of native wages on the X vector in year t. Then, we predict residuals for each immigrant and find the corresponding percentile θ_{it} in the residual distribution of natives. Using this method, we obtain three measures of relative wage depending on specification of the X vector: (i) unconditional if X includes only the intercept; (ii) age-specific if X also contains a quartic polynomial in age; and (iii) conditional if X includes the level of schooling, urban residence, and location in West Germany in addition to the intercept and a quartic polynomial in age.
- **Change in relative wage (percentile), $\dot{\theta}_{it}$,** is the average annual change in relative wage over the future 5-year period, $\dot{\theta}_{it} = \left(\frac{1}{5}\right)\sum_{t}^{t+4}(\theta_{it+1} - \theta_{it})$.

Employment participation

A dummy variable that takes the value of 1 if the respondent is working and the value of 0 if not-working. It is based on the labor force status from the dataset of generated variables (SOEP, 2014b).

Unemployment

A dummy variable that takes the value of 1 if the respondent is not-working and registered unemployed and the value of 0 if the respondent is working. It is based on the labor force status from the dataset of generated variables (SOEP, 2014b). The probability of unemployment is conditional on being in the labor force.

Average distance to work by region

We calculate the average distance (in kilometers) between home and workplace by state and year to proxy for fixed costs associated with work. The variable is constructed using individual reports on commuting distance from home to work available in PL file, which is then averaged at the state-year level. The distance is top coded at 200 km. The information is available for selected years and the values for missing years are taken from the neighboring year: 1984-87 from 1985, 1988-89 from 1990, 1991-92 from 1993 (and 1990 for East Germany), 1994 and 1996 from 1995, 1997 and 1999 from 1998, and 2000 from 2001. After 2000, the question on commuting distance is asked every year. Individuals who have workplace and home in the same building are assigned a zero distance. Individuals whose location of work varies or answered `difficult to say' are assigned a missing value for the distance.

Perceived discrimination because of the country of origin

Two binary variables are created from the survey question: "How often have you personally felt disadvantaged in Germany in the last two years because of your origins?" The first variable takes the value of 1 if the answer is "sometimes" and the value of 0 if the answer is "never". The second variable takes the value of 1 if the answer is "frequently" and the value of 0 if the answer is "never". The immigrant respondents were asked this question in 1996-2011 and 2013 surveys. We did not include answers of natives who are born in Germany without German citizenship. For consistency, we also had to exclude a newly drawn sample of immigrants in 2013 (Sample M) because the formulation of discrimination questions for this sample applies to the previous experience in general and does not refer to the last two years.

Remains in the panel

A binary indicator that takes up the value of one if a respondent in year *t* remains in the GSOEP panel after 5 years, and the value of zero if a respondent exits the survey during the following 5 years.

Mode of interview

The GSOEP uses several different modes of interviews, which we classify into 3 categories: [1] face-to-face, [2] self-written and mailed, and [3] computer assisted. Web-based interviews are combined with computer assisted personal interviews into one category.

Interviewer

All household interviews are classified into three categories: [1] first-time interview, [2] recurring interview with the same interviewer as in the previous round, and [3] recurring interview with a different interviewer compared to the previous round.

Supply of training offers and excess demand for training

The German Federal Ministry of Education and Research publishes annual statistics on the demand and supply of training contracts for the West and East Germany separately. Based on these statistics, we construct the log of supply of training offers and the excess demand for training. The excess demand is calculated as the difference between the log of unplaced training applicants and the log of unfilled training places still registered with employment offices (Federal Ministry of Education and Research, 2015).

B. Home Country Characteristics

GDP per capita

GDP numbers are taken from multiple sources. To make numbers consistent across sources, we first build an annual growth series for GDP per capita in constant prices. In 98 percent of our sample, we use the Conference Board Total Economy Database (TED, 2015), from which we extract the growth rate of PPP-adjusted GDP per capita in 1990 international dollars between 1960 and 2014. Missing values are replaced with real growth rates obtained from the Maddison Project (2013) and the World Development Indicators (WDI, 2016). The former source employs the same definition of GDP per capita as in TED (2015), while the latter source reports PPP-adjusted real GDP per capita in constant 2011 international dollars.[1]

For some countries that split apart (e.g., Czechoslovakia, Yugoslavia), the Maddison Project publishes the growth series for country parts before the breakup. However, GDP per capita is not available in any source for ex-USSR republics before 1980. Since some immigrants came to Germany from the former Soviet Union before 1980, we use real wage growth instead of GDP per capita growth for the Soviet republics between 1960 and 1980. Real wage growth is obtained from inflation-adjusted monthly wage series reported by the Central Statistical Board of the USSR.

The above four sources provide a complete time series on annual real growth of GDP per capita ($\dot{y}_{c[t-1,t]}$ in country c and year t compared to previous year) for all countries in GSOEP sample between 1961 and 2015. By using this growth series and the PPP-adjusted GDP per capita values in 2011 as a baseline (WDI, 2016), we construct a time-series of PPP-adjusted GDP per capita in constant 2011 international dollars, y_{ct}.

Using $\dot{y}_{c[t-1,t]}$ and y_{ct}, we construct the following variables:

- GDP per capita in the home country in the year of arrival, $y_{c,t=a}$;
- Average annual growth of GDP per capita in the home country over the next 5-year period, $\left(\frac{1}{5}\right)\sum_{t}^{t+4}\dot{y}_{c[t,t+1]}$;
- Average annual growth of GDP per capita in Germany over the next 5-year period.

[1] Simple coefficient of correlation between the TED and Maddison series of per capita GDP growth is 0.92 and between the TED and WDI series is 0.91.

Political instability

We capture political instability in a home country by using the dataset on Major Episodes of Political Violence (1946-2014) published by the Center for Systemic Peace (2015). This dataset assigns an integer score between 0 and 10 to each major episode of the war for independence, international violence/warfare, civil violence/warfare, and ethnic violence/warfare, where 0 indicates no episodes of political violence, 1 denotes sporadic political violence, and 10 stands for extermination and annihilation. All these scores are summed up into a combined index of political violence, which in our sample varies from 0 (74 percent of all immigrants) to 14 (Iraq in 1986). The original source does not provide scores for parts of former unified countries. Since many immigrants came from the former Soviet Union and ex-Yugoslavia, we use a variety of web sources to create the index of political violence for each republic before the breakup.

Based on the index of political violence, we construct two variables. The first variable is measured at the time of arrival to Germany. For easier interpretation, the index is aggregated into four distinct categories: 0=”no episodes of political violence, 1 or 2=“limited political violence”, 3=“serious political violence”, 4 and above=“warfare”. In the category of limited political violence, events are confined to short periods or specific areas; some population dislocation may occur; attributable deaths are up to ten thousand. Some examples from our sample include Czech Republic 1968, Turkey 1981-1983, Russia 1990, and China 1998. In the category of serious political violence, events are longer and involve a limited use of destruction technologies; population dislocations are in the tens of thousands of people; attributable deaths range from ten to fifty thousand. Examples include Syria 1973, Croatia 1992-1995, Tajikistan 1993-1995, and Kosovo 1996-1999. In the last category of warfare, events involve a broad use of destruction technologies and large dislocations of people; attributable deaths exceed 50,000 people. Examples include Afghanistan 1978-2001, Iran-Iraq 1980-1988, Armenia-Azerbaijan 1991-1994, Bosnia and Herzegovina 1992-1995, and Syria 2011 to present.

The second variable is the average political violence score in the country of origin during the time when an individual was 6 to 10 years old. To retain the sample size, we extend the index before 1946 for a small number of older immigrants who are born before the World War II (less than 1 percent of the estimation sample). Pre-1946 scores are assigned based on the detailed methodology provided by the Center for Systemic Peace (2015).

More developed country

A binary indicator for the above-the-median development index in the home country at the time of arrival. The composite development index is annual. It is constructed for the 1961-2014 period as a weighted average of four components: the log of PPP-adjusted real GDP per capita (+0.496), the index of political violence (-0.193), life expectancy at birth for both genders (+0.608), and the under-five mortality rate per 1,000 live births (-0.590). The weights in parentheses are obtained using the first principal component.

Neighboring country

A binary indicator that takes the value of one if the country of origin shares the border with Germany.

Linguistic proximity

The linguistic proximity between the primary language(s) of home countries and Standard German was calculated using the language trees classification provided by Ethnologue (2016). For the primary language, we chose either official language or the most spoken language in countries with multiple official languages. For example, we chose Hindi for India even though English is a second official language. The information on the number of people who speaks each language by country is also provided in Ethnologue (2016). If the country does not have a dominant language (e.g., there are two *equally* spoken languages), then the linguistic proximity is calculated for each language separately, and the final score is averaged (some examples include Chad, Cyprus, Kenya, Switzerland, etc.).

The variable takes five possible values based on the primary language's proximity to German in the language family tree: 0 for languages belonging to a separate family tree (e.g., Arabic, Turkish), 0.25 for languages that share the tree with German (e.g., French, Greek, Italian, Polish), 0.5 for languages that share the tree branch with German (e.g., Danish, Norwegian, Swedish), 0.75 that share the sub-branch (Afrikaans, Dutch, English), and 1 for German language. For instance, immigrants from Austria and Liechtenstein are assigned a linguistic proximity score of 1.

A binary indicator "same family language" is also created. It takes the value of one if the linguistic proximity is 0.25 or above.

Ethnic networks

We use the officially reported number of foreign population by country of origin as a percentage of the total German population. Foreign population consists of people who have the citizenship of their home country. It excludes naturalized German citizens, whom we consider as immigrants. Although this measure underestimates the share of foreign-born population, it captures well the major waves of migration from specific countries.

The home country share of foreign population can be obtained with sensible imputations from 1961 and onwards. The OECD International Migration Database (OECD, 2016a) and the German Central Register of Foreign Nationals (Ausländerzentralregisters) report the annual composition of foreign population by origin, which covers more than 99 percent of foreign population from 1990 and onwards, 91 to 96 percent from 1969 to 1989, 85 to 90 percent in 1967-1968, and 71 percent in 1961. Missing values for years 1962-1966 and occasional intermittent missing values in other years are imputed using a simple country-specific linear interpolation. Missing values for the home country share in population in the 1960s and 1970s are set to zero if the country has not achieved the 0.05 percent of total population in the 1980s. In the GSOEP estimation sample, 8 percent of *SMIGPOP* values are interpolated and 3 percent are set to zero.

We split our network measure into (i) established networks based on the stock of immigrants from the same country of origin 5 years ago and (ii) recent networks measured as additional flows of immigrants from the same country of origin during the last 5 years.

Time trend

Year of survey minus 1983.

D. References for data sources

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg, 2003. "Fractionalization," *Journal of Economic Growth* 8: 155-194.

Ausländerzentralregisters (Central Register of Foreign Nationals). Ausländische Bevölkerung (Foreign Population Statistics), Statistisches Bundesamt (Federal Statistical Office), annual.

Bollinger, Christopher R. and Barry T. Hirsch, 2006. "Match Bias from Earnings Imputation in the Current Population Survey: The Case of Imperfect Matching," *Journal of Labor Economics* 24(3): 483-519.

CEPII, 2016a. GeoDist Distance Measures, http://www.cepii.fr/

CEPII, 2016b. Language Dataset, http://www.cepii.fr/

Center for Systemic Peace, 2015. *Major Episodes of Political Violence, 1946-2014*, http://www.systemicpeace.org/inscrdata.html

Ethnologue, 2016. *Languages of the World*, 19th edition http://www.ethnologue.com

Federal Ministry of Education and Research, 2015. Report on Vocational Education and Training, annual.

Mayer, Thierry and Soledad Zignago, 2011. "Notes on CEPII's Distances Measures: The GeoDist Database," *CEPII Working Paper* 2011-25.

Melitz, Jacques and Farid Toubal, 2014. "Native Language, Spoken Language, Translation and Trade," *Journal of International Economics*, 92(2): 351-363.

OECD, 2016a. OECD International Migration Database, http://www.oecd.org/els/mig/keystat.htm

OECD, 2016b. OECD.Stat (database). http://stats.oecd.org/

SOEP, 2014a. *Documentation on Biography and Life History Data for SOEP v31 and v31.1*, edited by Jan Goebel. SOEP Survey Papers 312: Series D. Berlin: DIW/SOEP

SOEP, 2014b *Documentation of Person-related Status and Generated Variables in $PGEN for SOEP v31.1*. SOEP Survey Papers 307: Series D. Berlin: DIW Berlin/SOEP

The Maddison-Project, 2013 version, http://www.ggdc.net/maddison/maddison-project/home.htm

TED, 2015. The Conference Board Total Economy Database, May 2015, http://www.conference-board.org/data/economydatabase/

WDI, 2016. World Development Indicators, World Bank, http://data.worldbank.org/

Appendix A3. Comparative Statics

In this appendix, we derive the relationship between the wage convergence ($\dot{w}_m - \dot{w}_n$) and the change in skill price, $\varphi_m = (1 + \dot{p}_m)$ and $\varphi_n = (1 + \dot{p}_n)$.

From equation (10'') and assuming $\tau = 1$ for natives, we have

$$(\dot{w}_m - \dot{w}_n) \approx \left[([r\alpha(1+\dot{p}_m)]^{-1} + 1) \left[\frac{r\alpha A_m (1+\dot{p}_m) K^{\alpha+\beta-1}}{\tau} \right]^{\frac{1}{1-\alpha}} + \dot{p}_m \right] - \left[([r\alpha(1+\dot{p}_n)]^{-1} + 1) \left[r\alpha A_n (1+\dot{p}_n) K^{\alpha+\beta-1} \right]^{\frac{1}{1-\alpha}} + \dot{p}_n \right]$$

Let

$$f \approx (r\alpha)^{\frac{\alpha}{1-\alpha}} K^{\frac{\alpha+\beta-1}{1-\alpha}} \left[\left(\frac{A_m}{\tau}\right)^{\frac{1}{1-\alpha}} \left(\varphi_m^{\frac{\alpha}{1-\alpha}} + r\alpha \varphi_m^{\frac{1}{1-\alpha}} \right) - A_n^{\frac{1}{1-\alpha}} \left(\varphi_n^{\frac{\alpha}{1-\alpha}} + r\alpha\, \varphi_n^{\frac{1}{1-\alpha}} \right) \right] + \varphi_m - \varphi_n,$$

where $f = (\dot{w}_m - \dot{w}_n)$, $\varphi_m = (1 + \dot{p}_m)$ and $\varphi_n = (1 + \dot{p}_n)$.

$$\frac{\partial f}{\partial \varphi_m} = (r\alpha)^{\frac{\alpha}{1-\alpha}} K^{\frac{\alpha+\beta-1}{1-\alpha}} \left[\left(\frac{A_m}{\tau}\right)^{\frac{1}{1-\alpha}} \left(\frac{\alpha}{1-\alpha} \varphi_m^{\frac{2\alpha-1}{1-\alpha}} + \frac{r\alpha}{1-\alpha}\, \varphi_m^{\frac{\alpha}{1-\alpha}} \right) \right] + 1 > 0$$

$$\frac{\partial f}{\partial \varphi_n} = -(r\alpha)^{\frac{\alpha}{1-\alpha}} K^{\frac{\alpha+\beta-1}{1-\alpha}} \left[(A_n)^{\frac{1}{1-\alpha}} \left(\frac{\alpha}{1-\alpha} \varphi_n^{\frac{2\alpha-1}{1-\alpha}} + \frac{r\alpha}{1-\alpha}\, \varphi_n^{\frac{\alpha}{1-\alpha}} \right) \right] - 1 < 0$$

For $h = \frac{\varphi_n}{\varphi_m}$

$$\frac{\partial h}{\partial \varphi_n} = \frac{1}{\varphi_m} \qquad \text{(A4-1)}$$

$$\frac{\partial h}{\partial \varphi_m} = -\frac{\varphi_n}{\varphi_m^2} \qquad \text{(A4-2)}$$

Hence, the change in wage divergence with respect to the change in relative growth in skill prices of natives compared to immigrants is given by:

$$\frac{\Delta f}{\Delta h} = \frac{\frac{\partial f}{\partial \varphi_m}\Delta\varphi_m + \frac{\partial f}{\partial \varphi_n}\Delta\varphi_n}{\frac{\partial h}{\partial \varphi_m}\Delta\varphi_m + \frac{\partial h}{\partial \varphi_n}\Delta\varphi_n} \qquad \text{(A4-3)}$$

If we hold φ_m constant ($i.e.,\ \Delta\varphi_m = 0$), we get:

$$\frac{\Delta f}{\Delta h} = \frac{\frac{\partial f}{\partial \varphi_n}}{\frac{\partial h}{\partial \varphi_n}}$$

$$\frac{\Delta f}{\Delta h} = \frac{-(r\alpha)^{\frac{\alpha}{1-\alpha}} K^{\frac{\alpha+\beta-1}{1-\alpha}} \left[(A_n)^{\frac{1}{1-\alpha}} \left(\frac{\alpha}{1-\alpha} \varphi_n^{\frac{2\alpha-1}{1-\alpha}} + \frac{r\alpha}{1-\alpha} \varphi_n^{\frac{\alpha}{1-\alpha}} \right) \right] - 1}{\frac{1}{\varphi_m}} < 0$$

Similarly, if we hold φ_n constant ($i.e.,\ \Delta\varphi_n = 0$), we get:

$$\frac{\Delta f}{\Delta h} = \frac{\frac{\partial f}{\partial \varphi_m}}{\frac{\partial h}{\partial \varphi_m}}$$

$$\frac{\Delta f}{\Delta h} = \frac{(r\alpha)^{\frac{\alpha}{1-\alpha}} K^{\frac{\alpha+\beta-1}{1-\alpha}} \left[\left(\frac{A_m}{\tau}\right)^{\frac{1}{1-\alpha}} \left(\frac{\alpha}{1-\alpha} \varphi_m^{\frac{2\alpha-1}{1-\alpha}} + \frac{r\alpha}{1-\alpha} \varphi_m^{\frac{\alpha}{1-\alpha}} \right) \right] + 1}{-\frac{\varphi_n}{\varphi_m^2}} < 0$$

From equations (A4-1) and (A4-2), the gradient (direction of steepest ascent) of h is:

$$(\Delta\varphi_n,\ \Delta\varphi_m) = \left(\frac{1}{\varphi_m},\ -\frac{\varphi_n}{\varphi_m^2}\right)$$

Plugging these values into equation (A4-3), we can see that f also decreases along the gradient of h.

Appendix A3 Tables and Figures

Table W1: Wage Equation with Years Since Migration

	OLS	OLS	Mixed	Mixed
	(1)	(2)	(3)	(4)
Years since migration	0.011***	0.010***	0.010***	0.010***
	(0.001)	(0.001)	(0.001)	(0.001)
Age	0.250***	0.261***	0.328***	0.323***
	(0.069)	(0.063)	(0.065)	(0.065)
Age^2	-0.007***	-0.007***	-0.010***	-0.009***
	(0.003)	(0.002)	(0.002)	(0.002)
$Age^3 \times 10^3$	0.075*	0.077**	0.119***	0.116***
	(0.041)	(0.037)	(0.038)	(0.037)
$Age^4 \times 10^5$	-0.029	-0.029	-0.054**	-0.052**
	(0.024)	(0.021)	(0.022)	(0.022)
Female	-0.283***	-0.297***	-0.282***	-0.295***
	(0.007)	(0.006)	(0.010)	(0.010)
Adjusted years of schooling	0.046***	0.042***	0.040***	0.035***
	(0.002)	(0.002)	(0.003)	(0.003)
Urban residence	0.052***	0.061***	0.062***	0.067***
	(0.009)	(0.008)	(0.013)	(0.013)
Intercept	-1.709**	-1.823***	-2.457***	-2.496***
	(0.678)	(0.626)	(0.652)	(0.647)
Year FE	Yes	Yes	Yes	Yes
Home country FE	No	Yes	No	Yes
R^2	0.233	0.306	...	...
Standard deviation of $\hat{a}_i$, $\hat{\sigma}_a$	...	...	0.386	0.370
			(0.014)	(0.014)
Standard deviation of $\hat{b}_i$, $\hat{\sigma}_b$	...	...	0.017	0.017
			(0.001)	(0.001)
Correlation ($\hat{a}_i$, $\hat{b}_i$), $\hat{\rho}_{ab}$	...	...	-0.749	-0.764
			(0.022)	(0.022)

Notes: N=28,227. Table presents the estimates of wage equation with a random intercept a_i and a random slope b_i on years since migration. The dependent variable is the log of hourly wage rate. The estimates correspond to Equation (1) and use sampling weights. There are 31 year fixed effects and 119 home country fixed effects. Robust standard errors are in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table W2: Bivariate Selection Probit Model, Marginal Effects

Variables	***Stays in panel***	***Reports wage growth***	***Variables***	***Stays in panel***	***Reports wage growth***
Pre-migration years of			Instability in home country		
education	0.000	0.006***	at arrival		
	(0.001)	(0.001)	Limited political violence	-0.010	-0.030***
Female	0.029***	-0.242***		(0.008)	(0.010)
	(0.005)	(0.005)	Serious political violence	0.042***	-0.177***
Ethnic German	0.099***	0.025***		(0.012)	(0.014)
	(0.007)	(0.009)	Warfare	0.006	-0.080***
Age group				(0.010)	(0.011)
26-35	0.054***	0.093***	Ethnic networks	0.020***	-0.020***
	(0.012)	(0.014)		(0.003)	(0.004)
36-45	0.047***	0.144***	Average commuting		-0.008***
	(0.012)	(0.014)	distance		(0.002)
46-55	0.085***	0.067***	GDP per capita growth in	-0.008***	-0.002***
	(0.012)	(0.014)	home country [t+1, t+5]	(0.001)	(0.000)
55-65	0.036***	-0.320***	Mode of interview		
	(0.013)	(0.015)	Self-written and mailed	-0.143***	-0.038***
Parents' education				(0.011)	(0.006)
General sec and upper	-0.014*	-0.029***	Computer assisted	0.007	0.002
vocational	(0.008)	(0.010)		(0.008)	(0.002)
Higher education	0.031***	-0.072***	Interviewer		
	(0.012)	(0.014)	First interview	-0.081***	-0.021***
Urban residence	-0.004	0.018**		(0.012)	(0.005)
	(0.007)	(0.008)	Different interviewer	-0.041***	-0.011***
Log of GDP per capita in	-0.001	0.002		(0.008)	(0.003)
home country at arrival	(0.006)	(0.008)	Intercept	Yes	Yes
Linguistic proximity	0.106***	0.206***	Year FE	Yes	Yes
	(0.019)	(0.023)	N	37,253	24,708

Notes: Table presents the joint maximum likelihood estimates of two probit equations: one for staying in the survey in t+5 conditional on being interviewed in t (column 1) and another one for non-missing wage growth conditional on staying in the survey between t and t+5 (column 2). Reported are the marginal effects (MEs) evaluated at sample means. The means can be found in column "Sample B" of Table 3. Robust standard errors are in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Base/omitted categories are shown in notes to Table 3. The Wald LR test of independent equations of $corr(\upsilon_1, \upsilon_2) = 0$: chi-squared = 40.91***; $\widehat{corr}(\upsilon_1, \upsilon_2) = -0.49$.

Table W3: Auxiliary Equations for Estimating the Treatment Effect of Post-Migration Education on Immigrants' Wage Convergence

	Selection into Treatment	*Outcome model for ATE*		*Outcome model for ATT*	
		T=0	*T=1*	*T=0*	*T=1*
Pre-migration years of education, *a*	0.136***	0.049*	-0.051	0.058	0.074**
	(0.007)	(0.029)	(0.063)	(0.044)	(0.037)
Female	-0.050*	0.186*	0.777**	0.154	0.560***
	(0.029)	(0.112)	(0.385)	(0.190)	(0.217)
Ethnic German	0.578***	0.014	1.245***	-0.224	0.649**
	(0.042)	(0.173)	(0.384)	(0.259)	(0.276)
Age at *t*					
26-35	...	2.923***	1.157*	3.004***	0.283
		(0.550)	(0.616)	(0.580)	(0.511)
36-45	...	2.220***	1.342**	2.136***	0.783
		(0.545)	(0.675)	(0.581)	(0.526)
46-55	...	2.376***	1.321*	2.005***	-0.077
		(0.551)	(0.764)	(0.605)	(0.559)
55-65	...	3.801***	2.725**	3.546***	0.646
		(0.592)	(1.123)	(0.687)	(0.706)
Age at *a*	-0.083***	...	...	...	...
	(0.003)				
Parents' education					
General sec and upper vocational	0.481***	-0.142	0.073	-0.210	-0.165
	(0.044)	(0.191)	(0.359)	(0.233)	(0.268)
Higher education	0.592***	0.495	1.068**	0.333	1.147***
	(0.060)	(0.410)	(0.464)	(0.525)	(0.399)
Urban residence at *t*	-0.222***	0.142	-0.658*	0.247	-1.558***
	(0.039)	(0.153)	(0.339)	(0.253)	(0.310)
Log of GDP per capita in home country at *a*	-0.018	0.639***	0.259	0.655***	0.351
	(0.037)	(0.170)	(0.354)	(0.214)	(0.270)
Instability in home country at *a*					
Limited political violence	0.133***	-0.827***	1.851***	-0.879***	0.655*
	(0.049)	(0.187)	(0.622)	(0.263)	(0.347)
Serious political violence	0.148*	-0.241	2.246**	-0.184	1.394**
	(0.084)	(0.415)	(1.062)	(0.586)	(0.703)
Warfare	0.458***	0.616**	0.906*	0.328	0.433
	(0.058)	(0.259)	(0.473)	(0.315)	(0.398)
Ethnic networks at *t*	...	0.282***	0.099	0.323**	0.317*
		(0.080)	(0.244)	(0.132)	(0.187)
Ethnic networks at *a*	-0.206***				
	(0.024)				
Linguistic proximity	-0.669***	-1.026**	0.539	-0.413	-0.634
	(0.102)	(0.440)	(0.879)	(0.601)	(0.653)
Log of supply of training offers at *a*	1.178***	...	...	...	...
	(0.140)				
Excess demand for training at *a*	0.083***	...	...	...	...
	(0.020)				
Political violence score	-0.068***	...	...	...	...

in home country, ages 6-10	(0.012)				
Linear time trend	-0.002	0.133***	0.071	0.094*	0.022
	(0.008)	(0.031)	(0.096)	(0.048)	(0.068)
Linear time trend squared	0.001***	-0.005***	-0.004	-0.003*	-0.000
	(0.000)	(0.001)	(0.003)	(0.002)	(0.002)
Intercept	-15.560***	-10.005***	-3.603	-10.107***	-4.043
	(1.741)	(1.729)	(3.379)	(2.158)	(2.610)

Notes: N=13, 144. Table presents the estimates of auxiliary equations, which are used in calculating the average treatment effects of post-migration education on immigrants' wage convergence. Both ATE and ATT from these estimates are reported in Table 8, line 1. Robust standard errors are in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Treatment effects are obtained using doubly robust inverse-probability weighted regression adjustment. Treatment (T) is investment in any post-migration education in Germany. The selection model is estimated using probit with all covariates taken at the beginning of the investment period, that is, at arrival. The outcome model is the OLS model of wage convergence shown in Table 4, except that it is split into two subsamples (T=1 and T=0). Base/omitted categories are shown in notes to Table 3. About 200 observations with the propensity score outside the range [0.001, 0.999] are dropped to satisfy the overlap conditions.

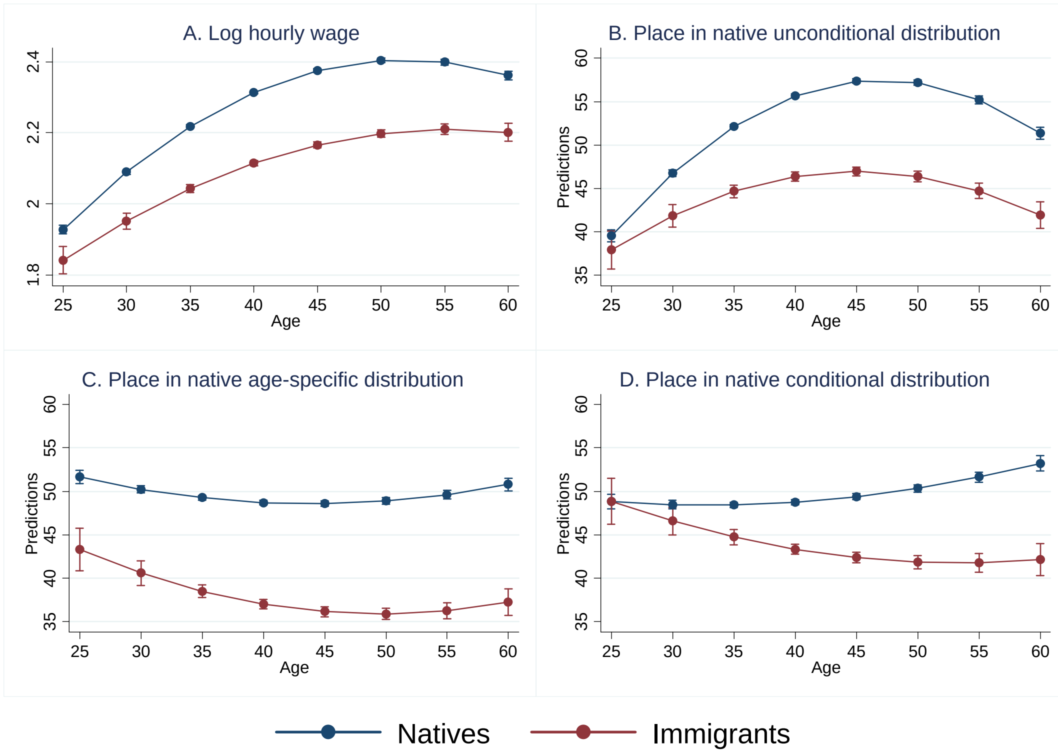
Table W4: OLS and IV Estimates of the Immigrants' Wage Convergence Model

	OLS	*Probit*	*2SLS*	*GMM*	*LIML*
Post-migration years of education	0.276**	...	3.241**	3.044*	3.750**
	(0.129)		(1.561)	(1.554)	(1.848)
Pre-migration years of education	0.048**	0.044***	-0.011	-0.011	-0.021
	(0.021)	(0.007)	(0.039)	(0.038)	(0.044)
Female	0.257***	-0.168***	0.302***	0.304***	0.309***
	(0.097)	(0.030)	(0.102)	(0.102)	(0.104)
Ethnic German	0.331**	0.477***	-0.019	0.001	-0.079
	(0.137)	(0.040)	(0.226)	(0.225)	(0.254)
Age at *t*					
26-35	1.544***	-0.640***	2.224***	2.121***	2.340***
	(0.383)	(0.070)	(0.549)	(0.547)	(0.598)
36-45	1.237***	-1.029***	2.344***	2.200***	2.534***
	(0.382)	(0.070)	(0.723)	(0.720)	(0.816)
46-55	1.257***	-1.268***	2.581***	2.410***	2.808***
	(0.389)	(0.075)	(0.826)	(0.822)	(0.941)
55-65	2.521***	-1.466***	3.953***	3.738***	4.199***
	(0.433)	(0.098)	(0.897)	(0.892)	(1.021)
Parents' education					
General sec and upper	-0.198	0.370***	-0.619**	-0.619**	-0.692**
vocational	(0.156)	(0.045)	(0.278)	(0.277)	(0.313)
Higher education	0.917***	0.065	0.409	0.367	0.321
	(0.281)	(0.062)	(0.401)	(0.399)	(0.438)
Urban residence	-0.294**	-0.157***	-0.204	-0.203	-0.189
	(0.137)	(0.041)	(0.144)	(0.143)	(0.147)
Log of GDP per capita in	0.406***	-0.136***	0.525***	0.506***	0.545***
home country at *a*	(0.146)	(0.038)	(0.159)	(0.158)	(0.164)
Instability in home country at *a*					
Limited political violence	-0.412**	0.009	-0.415**	-0.422***	-0.415**
	(0.162)	(0.051)	(0.162)	(0.162)	(0.163)
Serious political violence	0.161	0.010	0.547	0.441	0.613
	(0.342)	(0.082)	(0.399)	(0.397)	(0.421)
Warfare	0.539**	-0.010	0.317	0.364	0.279
	(0.219)	(0.061)	(0.246)	(0.244)	(0.258)
Ethnic networks at *t*	0.268***	-0.119***	0.332***	0.329***	0.343***
	(0.071)	(0.022)	(0.079)	(0.078)	(0.082)
Linguistic proximity	-0.884**	-0.029	-0.704*	-0.629*	-0.673*
	(0.359)	(0.105)	(0.378)	(0.375)	(0.386)
Log of supply of training offers at *a*	...	0.931***	...	...	...
		(0.128)			
Excess demand for training at *a*	...	0.089***	...	...	...
		(0.020)			
Political violence score	...	-0.124***	...	...	...
in home country, ages 6-10		(0.016)			
Linear time trend	0.124***	0.025***	0.107***	0.108***	0.104***
	(0.027)	(0.009)	(0.029)	(0.029)	(0.030)
Linear time trend squared	-0.004***	0.001**	-0.005***	-0.005***	-0.005***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)

Intercept	-6.494***	-11.876***	-8.501***	-8.188***	-8.846***
	(1.450)	(1.623)	(1.799)	(1.789)	(1.922)

Notes: N=13, 353. Table presents the OLS and IV estimates of wage convergence model. IV estimates are based on a 3-step procedure described in Section 4.4. The first step calculates the propensity score from the probit model of post-migration education reported in Column 2. 2SLS=two-stage least squares; GMM=generalized method of moments; LIML=limited-information maximum likelihood. Robust standard errors are in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Base/omitted categories are shown in notes to Table 3. Unknown parents' education is also included in the estimates but not shown here. The dependent variable is the annual change in immigrants' conditional relative wage averaged over the 5-year period.

Figure W1: Age-Wage Profiles, FE Estimates



Notes: The life-cycle profiles of the log of hourly wage and relative wages are predicted marginal effects from the FE regression of the corresponding outcome for a given group (natives or immigrants) on a quadratic polynomial in age with individual fixed effects and robust standard errors. The relative wage is defined as the placement of immigrants in the native wage distribution. Definitions of relative wage are discussed in Section 2. The 95 percent confidence interval for the point estimate is also shown.

Figure W2: Home Country Characteristics by Year of Migration

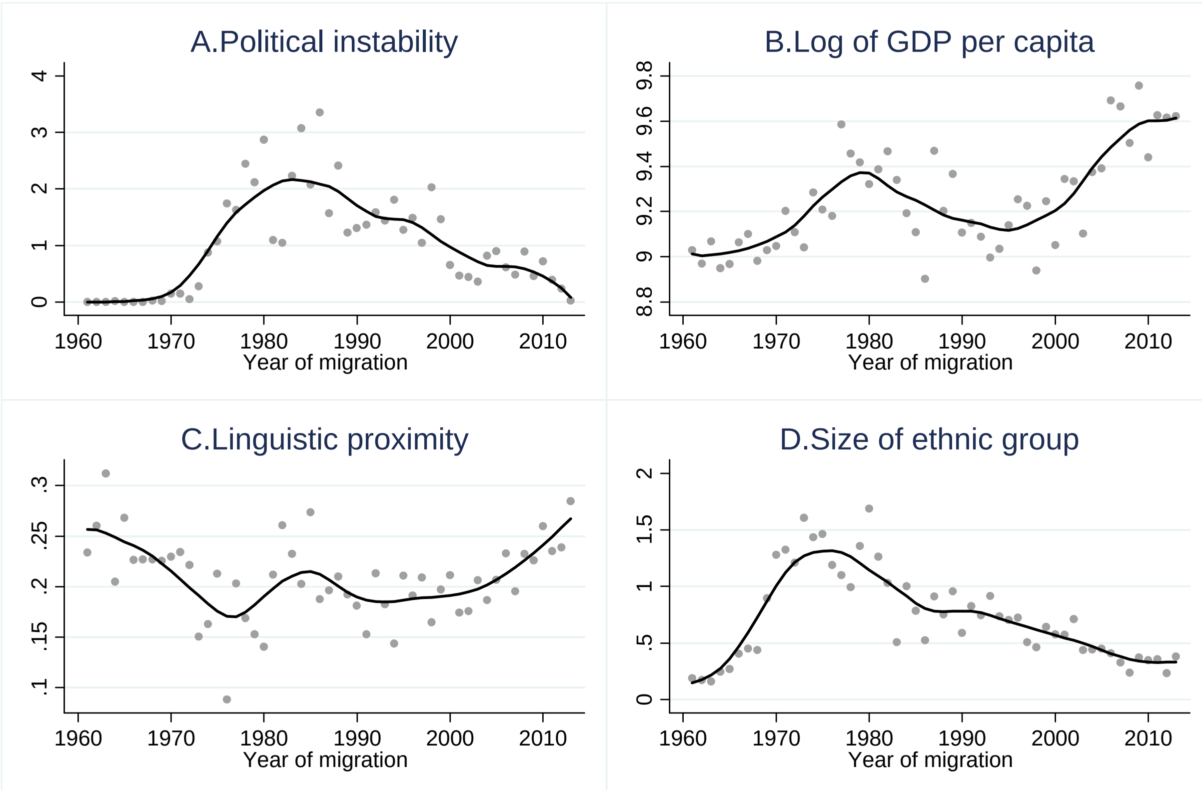
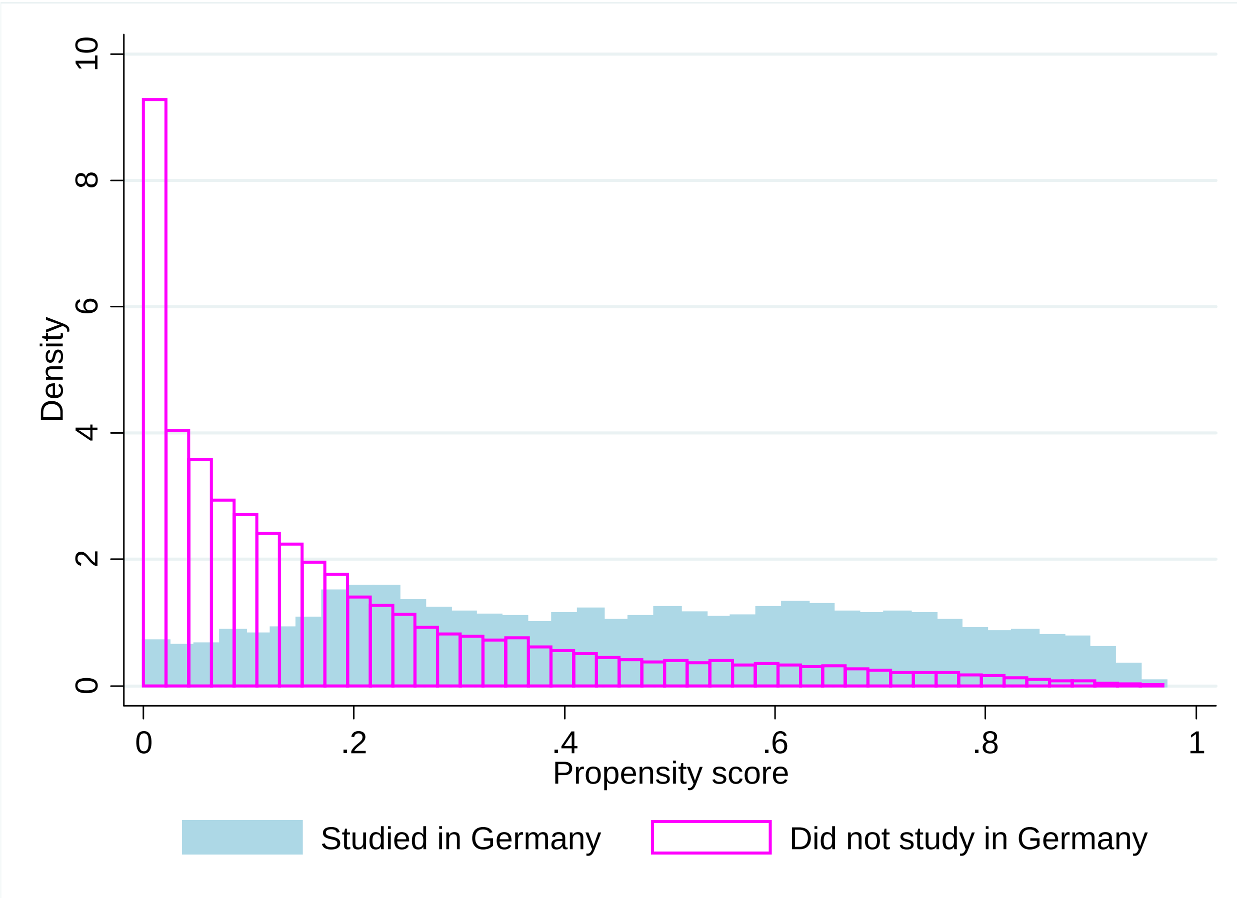


Figure W3: Common Support Region



Notes: Figure shows the distribution of the predicted probability of post-migration education for the two groups of immigrants: (i) those who studied in German schools or underwent job training after migration (treatment group) and (ii) those who did not receive any post-migration education (control group). The propensity score is predicted from the probit selection-into-treatment equation reported in Appendix Table W3.