University of Zurich[UZH]

Firm Heterogeneity and the Immigrant-Native Earnings Gap

Master's Thesis

Supervisor
Prof. Dr. Josef Zweimüller
Department of Economics
University of Zurich

Author:	Dominik Lukač
	dominik.lukac@econ.uzh.ch
	Basler Strasse 25/2
	79713 Bad Säckingen
Student ID:	21-714-373
Field of Study:	Economics

Date of Submission: March 27, 2024

Firm Heterogeneity and the Immigrant-Native Earnings Gap*

Dominik Lukač†

March 27, 2024

Abstract

Using rich administrative and survey data I measure the role of heterogeneous firm wage policies on the level and change of the immigrant-native earnings gap in Switzerland between 2002–2020. Firm heterogeneity is a major driver of both, variation in earnings and the immigrant-native earnings gap. Contrary to much of prior research both, *within* and *between* firm wage effects play an important role. Using the cohort of immigrants arrived between 2000–2004, I estimate that climbing the 'firm ladder' is an important driver of convergence in earnings. Moving to high premia firms contributes most of the effect, while unequal within-firm wage policies are utterly persistent. Compositional effects such as selective emigration do not play a role in this measurement. The ascent on the firm ladder is driven by a combination of higher job mobility after arrival and larger steps on the firm ladder. Detailed data allows for estimating the decompositions by origin × education, age-at-arrival and cohort subgroups revealing refined patterns not evident in the pooled sample.

*I thank Josef Zweimüller, David Card, Rafael Lalive and Fabrizio Zilibotti for their invaluable feedback for this thesis and Damian Osterwalder for important guidance, discussions and for being able to learn from him in a professional and personal sense. I thank the Hans-Böckler-Foundation for financial support during my studies.

†University of Zurich, dominik.lukac@econ.uzh.ch

Contents

List of Figures

List of Tables

List of Abbreviations

2SLS	Two-stage Least Squares
AKM	Abowd, Kramarz, and Margolis (1999)
BLM	Bonhomme, Lamadon, and Manresa (2019) Model
CBWs	Cross-border workers
CHK	Card, Heining, and Kline (2013)
CRE	Correlated Random Effects Model (Bonhomme et al., 2023)
ESS	Earnings Structure Survey
KSS	Kline, Saggio, and Sølvsten (2020)
OASI	Old age and survivors' insurance
OLS	Ordinary Least Squares
ppts	percentage points
SE	Structural Survey
STATPOP	Population and Households Statistics
ZEMIS	Zentrales Migrationsinformationssystem

1 Introduction

Immigration has always been a 'hot topic' in the political and social discourse no matter the context. Recent years have seen an unprecedented growth in international migration which has undoubtedly contributed to the surge of right-wing political parties and ideologies in many developed economies. In the light of the explosive political force of the matter, it is important to understand how, when and through which channels immigrants integrate to local labor markets.

Since the seminal paper of Abowd, Kramarz, and Margolis (1999) (hereafter AKM), vast evidence has shown that heterogeneous firm wage policies explain a substantive part of earnings dispersion. A recent strand of the literature, namely Damas de Matos (2017), Dostie et al. (2021) and Arellano-Bover and San (2024), has merged research on immigrant-native earnings gaps with empirical models identifying firm-specific wage premia à la AKM. This makes it possible to investigate to what extent the earnings gap can be explained by firm heterogeneity and how climbing the firm ladder towards high-wage firms contributes to earnings assimilation.[1]

However, previous studies suffered from a lack of monthly earnings data, no observable education or no exact immigration date leading to potential bias due to the changing composition of immigrants. This thesis tackles these drawbacks with rich administrative data linking social security records, the population register and the immigration register. Compared to prior research in this topic, the Swiss data are unique in including an education variable, monthly earnings, detailed immigrant origin information, a long time horizon and the exact date of immigration. This makes it possible to mitigate bias due to compositional effects. Using these advantages I aim to answer two main research questions. First, I will investigate to what extent firm heterogeneity contributes to the immigrant-native earnings gap and if the contribution depends on immigrants' origin or education. Related to this is the question of how important climbing the firm ladder is for earnings assimilation over time. Second, I analyze the mechanisms behind earnings assimilation through the firm ladder channel.

The Swiss labor market is particularly suited to analyze immigrant behavior. Switzerland is among the countries with the highest foreign population share and has experienced a massive influx of foreign labor since Switzerland and the European Union signed the 'Agreement on the Free Movement of Persons' in 1999 (Dorn and Zweimüller, 2021; Beerli, Indergand, and Kunz, 2023).

To address the research questions, I follow the literature and estimate separate AKM-type models for immigrants and natives using an annual population-level data set covering

[1]Throughout this thesis the word 'assimilation' will also be used to describe immigrants' improvement in an outcome when this outcome is already superior to that of comparable natives for some subgroups.

2002–2020. Standard diagnostics tests widely applied in the literature provide evidence in favor of the model's assumptions and reject strong job-match effects or bias due to endogenous mobility. The model decomposes log earnings into time-invariant person effects, a function of time-varying worker characteristics and firm-specific earnings premia. Using these wage components I estimate that, corrected for estimation error-induced bias, heterogeneous firm wage policies explain 7.4–8.7% of the variation in earnings, consistent with findings for other developed countries (e.g. Bonhomme et al. (2023)). Worker-firm sorting, i.e. the covariance of person and firm effects, is more pronounced for immigrants where it explains 14.9% of the variation in earnings while only explaining 6.9% for natives, consistent with Dostie et al. (2021) and Arellano-Bover and San (2024).

To answer the first research question, I calculate the difference of the expectations of the two regression equations, similar to e.g. Card, Cardoso, and Kline (2016) and Dostie et al. (2021), making it possible to decompose the immigrant-native earnings gap into components driven by compositional effects and firm wage policies. Using an Oaxaca (1973)-type decomposition the firm components can be further split into an effect driven by differential *sorting* of immigrants and natives to high-premium firms and an effect driven by differential *pay-setting* policies for natives and immigrants within the same firm.
Firm wage policies are an important driver of the earnings gap in the cross-section, explaining 22.1% of the overall difference in log earnings. This is mainly driven by differential pay-setting, 16.8%, while differential sorting explains 5.3% of the overall firm effect. This contradicts standard AKM models assuming constant firm effects across groups. Most of the existing literature on group-specific AKM models finds a more prominent role of sorting and a weak pay-setting effect. The importance of differential sorting increases for younger demographics.
Analyzing the earnings gap by education levels reveals rich patterns not evident in comparable studies. Although tertiary-educated immigrants allocate much better to high-paying firms, this effect is almost entirely offset by unfavorable immigrant-native pay-setting regimes, resulting in almost the same mean firm premia as tertiary-educated natives. Non-tertiary-educated immigrants earn lower mean firm premia than comparable natives, this is almost equally driven by sorting and pay-setting effects.
Looking at the same decomposition by gender, reveals that firm-specific wage policies explain 6.1 and 10.8% of the unadjusted gender wage gap for natives and immigrants respectively. This is less than estimated by e.g. Card, Cardoso, and Kline (2016) for Portuguese data.
Using a decomposition inspired by Haltiwanger (1997) I estimates that firm entry and exit to and from the labor market contribute to a decrease of the cross-sectional earnings gap both through allocating immigrants to high-premium firms and more equitable firm wage policies.

Related to the first research question is the question to which extent earnings assimilation over time is driven by climbing the firm ladder. To isolate assimilation paths from the influence of newly arriving immigrants (Borjas, 1985; Lubotsky, 2007) I restrict the immigrant sample to those arriving in 2000–2004. This is made possible by the detailed immigration information available. Afterwards, I estimate detailed earnings and firm premia gaps by time since arrival similar to Arellano-Bover and San (2024). Due to the high quality data I can do this for origin × education subgroups. Firm premia at arrival depend heavily on origin and education, with tertiary-educated immigrants and those from economically well developed countries being favored. This is mainly driven by these subgroups allocating to high-premium firms better. Growth in firm premia is incredibly important for earnings assimilation, being responsible for 31–38% of total growth. Well-educated immigrants reach the natives' firm premia level after 6–29 years depending on their origin, while non-tertiary-educated immigrants do not close the gap to their native peers within the analyzed time frame.
Growth in firm effects is almost entirely driven by immigrants sorting to high-premium firms, while pay-setting stays relatively constant over time since arrival. By restricting the cohort to those immigrants still active in 2020 similar to Abramitzky, Boustan, and Eriksson (2014) and Dustmann and Görlach (2015) I ensure these effects are not driven by selective emigration of underachieving individuals.
Immigrants' skill is highly correlated to their age at arrival. Immigrants aged 10–29 show adverse earnings and firm premia which can be explained by worse education accumulation, consistent with other literature (e.g. Schaafsma and Sweetman (2001); Alexander and Ward (2018)). This could either be driven by self-selection of these immigrants, i.e. individuals will not migrate during their education, or their human capital accumulation being distorted due to the change of environment in this crucial age for development. Consistent with this hypothesis, immigrants coming as children show more similar outcomes to natives.
Next, I estimate the same assimilation paths for cohorts arriving after 2004. Later cohorts have, on average, higher firm premia on arrival, almost entirely explained by better sorting. Assimilation paths are roughly parallel. These estimates are not driven by compositional effects.

To get behind the drivers of firm premia growth and answer the second research question, I estimate assimilation trajectories of other outcomes. In a model of constant firm premia, climbing the firm ladder is a function of the probability of an employer switch and the expected size of the step conditional on a change. Job mobility depends heavily on education. Non-tertiary immigrants are 70–90% more likely to change their employer in the first year after immigration relative to natives. This excess probability flattens almost entirely within 10 years after arrival. Tertiary-educated immigrants are initially less likely to change employers, but have elevated probabilities in the following years. Immigrants' steps on the firm ladder conditional on an employer change are 2.5–4.5 times larger than those of natives immediately after arrival, this fades out quickly. Although immigrant men are more mobile than women, females make larger steps on the firm ladder conditional on a change (compared to natives of the same gender). Even after 20 years, immigrants are clustered at firms, having almost double the share of immigrant coworkers compared to natives.

Contribution to Literature. This thesis contributes to multiple strands of the literature. First, it contributes to the extensive literature about the influence of firm heterogeneity on wage inequality (e.g. Abowd, Kramarz, and Margolis (1999); Card, Heining, and Kline (2013); Card, Cardoso, and Kline (2016); Card et al. (2018); Song et al. (2019)) and its subfield of analyzing earnings differences between groups using AKM effects (e.g. Card, Cardoso, and Kline (2016); Sorkin (2017); Gerard et al. (2021)), especially between natives and immigrants (Damas de Matos (2017); Dostie et al. (2021); Arellano-Bover and San (2024)). I show that the AKM model's assumptions hold on the Swiss labor market. The high pay-setting effect I find contradicts standard AKM models assuming constant firm premia across immigrants and natives, similar to Arellano-Bover and San (2024). Furthermore, this thesis contributes to the research about estimation error corrected variance decompositions using AKM models initially sparked by Andrews et al. (2008) by estimating the decomposition using methods developed by Kline, Saggio, and Sølvsten (2020) (hereafter KSS), Bonhomme, Lamadon, and Manresa (2019) and Bonhomme et al. (2023) for Switzerland.

Second, I contribute to the literature analyzing the assimilation of immigrants in their destination labor markets (e.g. Chiswick (1978); Borjas (1985); Lubotsky (2007); Abramitzky, Boustan, and Eriksson (2014); Dustmann and Görlach (2015); Rho and Sanders (2021)). In this literature heterogeneous firm wage policies have been a relatively unexplored channel of immigrant assimilation. Although e.g. Damas de Matos (2017) and Dostie et al. (2021) analyze this, Damas de Matos (2017) does not estimate separate firm effects while Dostie et al. (2021) do not observe the exact immigration date, both use relatively short panels of 8 respectively 9 years. Arellano-Bover and San (2024) combine exact time since arrival with a decomposition into differential sorting and pay-setting for the special setting of former Soviet Union-Jews immigrating to Israel. I contribute to this literature by estimating detailed cohort $\times$ origin $\times$ education assimilation trajectories for firm premia using a long panel, revealing important nuances not evident before.

Third, I contribute to the growing literature on the link between monopsonistic labor markets and immigration (e.g. Malchow-Møller, Munch, and Skaksen (2012); Naidu, Nyarko,

and Wang (2016); Amior and Stuhler (2022); Amior and Manning (2023); Dustmann, Ku, and Surovtseva (2024)). The extent of differential pay-setting gives an important benchmark about the ability of firms to wage-discriminate. The important and persistent pay-setting effect within the same education level measured in the Swiss data hints towards the ability of firms to discriminate between natives and immigrants of the same skill level. These insights can be used in the theory.
Last, I contribute to the literature on the Swiss labor market[2], immigration into it (e.g. Beerli, Indergand, and Kunz (2023)) and the role of firms (e.g. Winter-Ebmer and Zweimüller (1999); Beerli et al. (2021)).

Outline. The remainder of this thesis is structured as follows. Section 2 begins by deriving a model of monopsonistic wage determination which forms the theoretical foundation for the empirical AKM model and explains key concepts used for the identification of the model. Afterwards, section 3 describes the Swiss data and important peculiarities of the Swiss labor market. Section 4 describes the AKM framework and its assumptions. In addition, the empirical strategy for measuring the immigrant-native earnings gap in the cross-section and by time since arrival is outlined. Next, section 5 starts with discussing standard AKM diagnostics tests. Subsequently, results of the variance decomposition of log earnings and assortative matching are presented. Thereafter, in section 6 the immigrant-native earnings gap is decomposed in the cross-section as well as by time since arrival. Section 7 explores job mobility and other mechanisms driving the firm-ladder ascent, while section 8 concludes.

2 Theory – Monopsonistic Wage Determination

In this section a simple static monopsonistic wage-setting model based on Card et al. (2018) and Gerard et al. (2021) is derived and its implications for firm policies are summarized, setting the theoretical basis for the empirical approach used in section 4. In the model a large number of firms compete for workers with random idiosyncratic tastes for each employer. These random taste shocks equip employers with monopsonistic market power. Following the literature on monopsonistic wage-setting models, firms post group-specific offers and are not able to negotiate individually with workers.[3]

[2]See Lalive and Lehmann (2020) for an overview of the current state.
[3]See Card (2022) for a review of the literature on monopsonistic labor markets.

Labor Supply. Let $\mathcal{J}$ be the set of active firms employing two groups of workers $g \in \{N, M\}$. Each firm $j \in \mathcal{J}$ posts a pair of wages (w_{Nj}, w_{Mj}) which can be costlessly observed by the workers. Workers' utility function of being employed at firm j is given by multinomial logit style preferences:

$$U_{igj} = \delta_g^0 \ln(w_{gj} - b_g) + a_{gj}^0 + v_{igj}, \tag{1}$$

where b_g is the reference wage based on the value of non-employment, a_{gj}^0 is a firm- and group-specific non-pecuniary amenity, $\delta_g^0 > 0$ is a factor expressing the relative valuation of excess wage over amenities and v_{igj} is a idiosyncratic taste parameter the worker assigns to the job, which can arise from non-pecuniary match factors such as distance to work or interactions with coworkers and supervisors. Assume that $v_{igj} = \tau_g \varepsilon_{igj}$ where ε_{igj} follows a type-I extreme value distribution and $\tau_g > 0$ governs the variance of idiosyncratic preferences and determines the importance of idiosyncratic preferences relative to excess wage and amenities for group g. Dividing by τ_g, a monotone transformation that does not affect the preference structure, one can define $\delta_g \equiv \delta_g^0/\tau_g$ and $a_{gj} \equiv a_{gj}^0/\tau_g$. Firms are willing to hire any worker who is willing to work at the posted wage, hence by McFadden et al. (1973) the probability of choosing firm j for any single worker equals the probability of worker i's utility being maximized by firm j's wage posting:

$$p_{gj} \equiv P(\arg\max_{k \in \mathcal{J}}\{U_{igj}\} = j) = \frac{\exp(\delta_g \ln(w_{gj} - b_g) + a_{gj})}{\sum_{j \in \mathcal{J}} \exp(\delta_g \ln(w_{gj} - b_g) + a_{gj})} \text{ for } g \in \{N, M\}. \tag{2}$$

Assuming the number of firms to be large eliminates strategic interactions due to the inability of any firm to influence $\sum_{j \in \mathcal{J}} \exp(\delta_g \ln(w_{gj} - b_g) + a_{gj})$ significantly. Therefore, the above probability can be closely approximated using exponential probabilities:

$$p_{gj} \approx D_g \exp(\delta_g \ln(w_{gj} - b_g) + a_{gj}) \text{ for } g \in \{N, M\}, \tag{3}$$

where D_g is a group-specific constant common across firms. Assuming there are L_g workers, firm-specific labor supply can be written as

$$\ln L_{gj}(w_{gj}) = d_g + \delta_g \ln(w_{gj} - b_g) + a_{gj} \text{ for } g \in \{N, M\}, \tag{4}$$

where $L_{gj} = p_{gj} L_g$ and $d_g = \ln(D_g L_g)$. Firm-specific labor supply is upward-sloping, leading to market power on the employer side in this monopsonistic competition framework.[4] Using the above result, the labor supply elasticity $e_{gj} = \frac{\partial L_{gj}}{L_{gj}} / \frac{\partial w_{gj}}{w_{gj}}$ can be calculated as:

[4] As noted by Amior and Manning (2023) an upward-sloping labor supply can alternatively be justified by search frictions instead of idiosyncratic preferences.

$$e_{gj} = \frac{\delta_g w_{gj}}{w_{gj} - b_g} \text{ for } g \in \{N, M\} \tag{5}$$

As $\delta_g \to \infty$ the labor supply gets perfectly elastic, i.e. labor supply becomes perfectly elastic because excess wage is infinitely more valued than amenities and idiosyncratic taste, resulting in a competitive labor market with exogenous wage b_g. Therefore δ_g is referred to as the labor supply elasticity parameter hereafter.

Labor Demand. Assume that firms have production functions of the form

$$Y_j = T_j f(L_{Nj}, L_{Mj}), \tag{6}$$

where T_j is a firm-specific productivity shifter. Firms observe labor supply but not the workers' idiosyncratic preference draw and post cost-minimizing wage pairs which solve the following optimization problem:

$$\min_{w_{Nj}, w_{Mj}} w_{Nj} L_{Nj}(w_{Nj}) + w_{Mj} L_{Mj}(w_{Mj}) \text{ s.t. } T_j f(L_{Nj}, L_{Mj}) \geq Y_0. \tag{7}$$

The first order conditions can be expressed as

$$w_{gj} = \frac{e_{gj}}{1 + e_{gj}} T_j \frac{\partial f(\cdot)}{\partial L_{gj}} \lambda_j \text{ for } g \in \{N, M\}, \tag{8}$$

where λ_j is the Lagrange multiplier which equals the marginal cost of production at optimal production Y_0 and $T_j \frac{\partial f(\cdot)}{\partial L_{gj}}$ is the marginal product with respect to L_{gj} for $g \in \{N, M\}$. To maximize profits the firm will equate marginal costs with marginal revenue. Thus the term $T_j \frac{\partial f(\cdot)}{\partial L_{gj}} \lambda_j$ is the marginal revenue product of firm j. According to equation 8, optimal group-specific wages will be determined by the marginal revenue product of the respective group subtracting a markdown depending on the group's labor supply elasticity. The more inelastic the labor supply, the higher the markdown will be. Notice that labor supply becomes more elastic as the offered wage approaches the reference wage (see equation 5). This in turn leads to a low markdown, implying that firms with low wage offers will pay wages closer to the marginal revenue product of the worker.
Using the above expression, or by equating marginal factor costs with the marginal revenue product, wages from equation 8 can be rewritten as:

$$w_{gj} = \frac{1}{1 + \delta_g} b_g + \frac{\delta_g}{1 + \delta_g} T_j \frac{\partial f(\cdot)}{\partial L_{gj}} \lambda_j \text{ for } g \in \{N, M\}. \tag{9}$$

The optimal wage for a group is a weighted average between the group's reference wage and its marginal revenue product. The weighting depends on δ_g which in turn depends on the group's relative valuation of excess wage and the extent of variation in worker-specific preferences τ_g. The firm-specific amenities a_{gj} do not influence the optimal wages, rather they only shift the intercept of the labor supply curve of the respective firm through equation 4, governing the size of the firm depending on the production technology.

Linear Production. To proceed further, assumptions concerning the firms' production function and market power on the output good market have to be made. To keep things simple let the firms be price takers on the output good market, selling for the price P_j^0, and the production function be linear of the form

$$f(L_{Nj}, L_{Mj}) = L_j \equiv \theta_N L_{Nj} + \theta_M L_{Mj}, \tag{10}$$

where θ_g gives the efficiency units per worker in group g and L_j is a linear aggregator for the total efficiency units available to firm j. This implies perfect substitutability between workers of the groups. Using these, the optimal wages in 9 can be expressed as:

$$w_{gj} = \frac{1}{1+\delta_g} b_g + \frac{\delta_g}{1+\delta_g} T_j \theta_g P_j^0 \text{ for } g \in \{N, M\}. \tag{11}$$

With a further assumption that reference wages are proportional to the groups' productivities, i.e. $b_g = b\theta_g$, it is possible to write the logarithm of the optimal wages as:

$$\ln w_{gj} = \ln \frac{\theta_g b}{1+\delta_g} + \ln(1 + \delta_g R_j) \text{ for } g \in \{N, M\}, \tag{12}$$

where $R_j \equiv \frac{T_j P_j^0}{b} = \frac{T_j \theta_g P_j^0}{b_g}$ is the ratio between the marginal revenue product of labor at firm j and the reference wage. By defining value added per standardized unit of labor as $\mu_j \equiv P_j^0 Y_j / L_j = P_j^0 T_j$ it is possible to rewrite $R_j = \mu_j / b$, so R_j is the ratio of value added per standardized unit of labor and the reference wage for a hypothetical worker with $\theta = 1$, i.e. one efficiency unit of labor. Therefore R_j is a standardized measure of productivity. Under the above assumptions log wages are additively separable with a term depending on the productivity of the worker and a firm-specific component depending the firm's productivity R_j and the group's preference parameter δ_g.

Implications for AKM-type Models. Firms with $R_j \approx 1$, i.e. value added per worker equals the outside option, the wage offers equal the group's marginal productivity:

$$\ln w_{gj} \approx \ln(T_j \theta_g P_j^0) \text{ for } g \in \{N, M\}, \tag{13}$$

This can intuitively be explained by the fact that these 'marginally efficient' firms offer wages $w_{gj} \approx \theta_g b = b_g$ (equation 12). The firms' labor supply elasticity 5 tends to infinity, eliminating the monopsonistic market power such that the firm offers wages equal to the marginal revenue product of the worker, revealing the workers' productivity. This is important, as identification of firm premia in the empirical model described in section 4 relies on a normalization using exactly this fact.

Assuming a low value of $\delta_g R_j$, the firm-specific component can be described using a Taylor approximation leading to the form:

$$\begin{aligned} \ln w_{Nj} &= \ln \frac{\theta_N b}{1+\delta_N} + \delta_N R_j \\ \ln w_{Mj} &= \ln \frac{\theta_M b}{1+\delta_M} + \delta_M R_j, \end{aligned} \tag{14}$$

which can be rewritten as:

$$\ln w_{gj} = \alpha_g + \psi_j^g, \tag{15}$$

where $\psi_j^g \equiv \delta_g R_j$ is a group-specific firm premium of firm j and $\alpha_g \equiv \ln \frac{\theta_g b}{1+\delta_g}$ is a group-specific constant. The group's reservation wage, although having a positive effect on the wage, is subsumed into the group-specific constant.

The model has some implications which can be verified empirically. As the labor supply elasticity parameter is given by $\delta_g = \delta_g^0/\tau_g$, groups with a higher valuation of excess wage relative to non-pecuniary amenities, i.e. higher δ_g^0, and less dispersion in their idiosyncratic taste for specific firms, i.e. lower τ_g, are expected to have higher firm premia. In addition, if one group has a higher δ_g, this group will face a steeper firm ladder. Also, the difference between ψ_j^N and ψ_j^M will be higher at more profitable firms. As $\psi_j^M/\psi_j^N = \delta_M/\delta_N$, an estimate for δ_M/δ_N, i.e. the relative size of the supply parameters, can be obtained by using the identity $\psi_j^M = (\delta_M/\delta_N)\psi_j^N$.

As summarized by Amior and Manning (2023) there are many reasons to expect that immigrants have a lower labor supply elasticity than natives, which would be modeled by $\delta_N > \delta_M$, assuming the same reservation wage. Language barriers, less access to social networks, poor information on labor market institutions and visa-related restrictions might lead to less efficient immigrant job search, implying a lower elasticity. On the other hand, immigrants might be more willing to relocate for a job or, especially in high-wage

industries or additionally consider offers from other countries, advocating for a higher elasticity of labor supply.
The theoretical implications will be addressed and discussed when relevant throughout the thesis.

3 Data

To get a full picture of immigration and firm dynamics in the Swiss economy, a rich data base is necessary. To this extent I combine three register data sources and two survey data sets provided by several administrative entities to construct a new, population-spanning annual panel data set of firms and workers. The analysis focuses on immigrants and natives working in Switzerland at some point between 2002 and 2020.
Section 3.1 describes the Swiss data, sources and advantages when compared to the data used in the existing literature. Section 3.2 describes the samples used for the baseline analysis in detail.

3.1 Sample Selection

Data sources. The main data source is the Old age and survivors' insurance (hereafter OASI) register provided by the Central Compensation Office. It contains information on labor income, nationality, employers and length of employment spells for *all* individuals paying Swiss social security contributions between 1981 and 2020. This includes virtually all self-employed individuals, immigrants and, most importantly, cross-border workers, which constitute an important part of the Swiss workforce.
I link individual accounts of the OASI register with data from the 'Population and Households Statistics' (hereafter STATPOP) provided by the Federal Statistical Office (hereafter FSO). STATPOP is based on federal, cantonal and municipal information, covers the years 2010–2020, and gives information such as age, gender, date of immigration and municipality of residence for individuals residing in Switzerland.
To get detailed information on immigrants, I use the 'Zentrales Migrationsinformationssystem' (hereafter ZEMIS). In this data set the State Secretariat for Migration collects detailed information such as age, gender, municipality of residence, workplace and residence permit for all individuals, including cross-border workers, without Swiss citizenship since 2002.
The above data are complemented by two survey data sets. First, the Structural Survey (hereafter SE) is a detailed repeated cross-sectional survey among $300,000$ random permanent residents conducted annually since 2010. Second, the Swiss Earnings Structure

Survey (hereafter ESS) is a survey conducted every second year since 2012 on the firm level, collecting information on the firm and its employees. Both surveys include information on the education of the individuals, as a result information on education is available for a majority of natives and immigrants.[5] Moreover, the ESS includes information on the firms' industry, proving useful in identifying the model's parameters as described in section 4.2.

Construction. Although Lachowska et al. (2023) show that the assumption of time-invariant firm AKM effects holds reasonably well in US data, this assumption should not be stretched. For this reason and for the reason of availability of detailed information about immigrants, the main sample spans the time frame 2002–2020.

For each individual I keep the person-year observation with the highest earnings per year, following the literature.[6] I calculate monthly earnings by dividing annual earnings by the length of the employment spell in months. All monetary variables are adjusted to 2020 Swiss francs (CHF) using the consumer price index.

The data do not contain information on working hours. To restrict attention to full-time equivalent workers, I use workers aged 25–60 and exclude all observations with earnings below the 15th percentile by year. The results are not sensible to using a different threshold or the 32.5% of the national average wage such as Bonhomme et al. (2023). Furthermore, appendix B presents the main results for a sample of males only. There are only minor changes compared to the baseline. To minimize the influence of outliers, I winsorize monthly earnings at CHF $100,000$.

Individuals are assigned to labor market regions according to their municipality of residence. These are defined by the FSO in 2000 based on the matrix of commuting flows between municipalities so that most of the residents simultaneously work in the same labor market region (Federal Statistical Office, 2000, 2022a).

Every individual born in a foreign country without Swiss citizenship since birth is defined as an immigrant. This corresponds to the FSO's definition of first generation immigrants, except for assigning individuals born abroad, with Swiss citizenship since birth, but with both parents born abroad to the native group (Federal Statistical Office, 2022b).

Immigration Date. STATPOP includes information on the date of immigration for all immigrants. Just as in Arellano-Bover and San (2024) this allows for inference on the *actual* time since arrival to Switzerland without relying on proxies typically used, such as

[5] I interpolate education using the nearest observation for each individual.

[6] If there are two observations with the same earnings, I use the observation with the longer firm tenure. If this still does not give a unique person-year data set, I choose randomly.

the first appearance in employment records (Dustmann, Ku, and Surovtseva, 2024) or the time of initial application for a Social Security Number (Rho and Sanders, 2021). Proxies based on administrative data will usually miss years of informal employment and will ignore the fact that some immigrants immigrate as children or teenagers to Switzerland. CBWs play an important role in the Swiss labor market, especially after the gradual removal of all immigration restrictions for workers from the European Union since the announcement of the reform in 1999 (Beerli et al., 2021).[7] To include them in the assimilation analysis in section 7 where observed immigration dates are necessary, I set their immigration date to the first appearance on the Swiss labor market. This makes sense as these individuals are unlikely to have gathered exposure to Switzerland before their first cross-border job.

Firms. Firm identifiers are assigned to institutional units by so-called compensation offices. Compensation offices are equipped with a certain degree of discretion when assigning these. In some cases, a certain firm identifier unites several legally independent units of a firm. In other cases, it is possible that a single firm reports its employees under different firm identifiers depending on the division within the firm, e.g. management vs. production. In the first case it can be argued that the different entities are represented by the same firm-wide wage policy. In the second case, this can even be of advantage if the firm has different wage policies by firm division. Even if this is not the case, the AKM estimator is still unbiased when treating the divisions as separate firms, although there is a potential loss in statistical efficiency (Card, Heining, and Kline, 2013).
There are several administrative reasons for compensation offices to reassign firm identifiers. As described above, treating a single firm as separate units does not bias the AKM estimates. Nevertheless, this can be a problem for identifying firm entry and exit. To deal with this and decrease estimation error, I reassign firm identifiers in a two-step approach which identifies firms based on year-to-year worker flows. Appendix E gives more information on the implementation of the algorithm.

Advantages. The combined data set has several advantages. The earnings data are not censored at the social security maximum such as, for example, the data in Card, Heining, and Kline (2013) (hereafter CHK). This allows for a reliable analysis of the immigrant-native gap for the top percentiles of the earnings distribution. Also, the data include private, as well as public sector employees. Other literature, e.g. Bonhomme et al. (2023) or CHK restrict to private sector employees due to the nature of the respective data

[7] Beerli et al. (2021) analyze the effects of the abolition of labor market restrictions on Swiss firms and workers in the border regions.

sources, ignoring a significant part of the workforce. Furthermore, the Swiss data include self-employed workers.

As mentioned by Song et al. (2019), using annual earnings as the outcome variable (e.g. Song et al. (2019) or Dostie et al. (2021)) makes it impossible to distinguish between labor supply decisions and changes in earnings per unit of time. Therefore, estimated AKM effects may capture systematic differences in labor supply. This study uses monthly income as the main outcome measure, which rules out between-month labor supply effects, differences within a month, for example differences in weekly hours could still pose a problem. Detailed data on country of origin and education level allow analysing the immigrant-native earnings gap by origin × education subgroups to identify heterogeneous effects and assimilation paths. Furthermore, the availability of data on natives' education allows for a comparison of immigrants to their native peers of the *same* education level. Both have not been implemented in comparable literature yet (e.g. Dostie et al. (2021); Arellano-Bover and San (2024)). Additionally, the availability of exact immigration dates makes estimation of heterogeneous effects by e.g. cohorts or age-at-arrival possible.

Compared to existing literature, the Swiss data are especially interesting for analyzing differences between natives and immigrants. This is because immigrants make up a large part of the Swiss workforce. In fact, behind Luxembourg, Switzerland has the highest share of foreign nationals in their domestic population among European Economic Area members (Dorn and Zweimüller, 2021).[8] Immigrant observations make up roughly 36% of the full sample, compared to 18% in Dostie et al. (2021) or 19% in Arellano-Bover and San (2024), while simultaneously keeping a larger or similar sample size.

3.2 Summary Statistics

Table 1 provides a descriptive overview of the samples used for the main analysis. Columns (1) and (2) shows characteristics of the full sample for immigrants and natives working in Switzerland between 2002–2020. For the AKM model to be identified it is necessary to restrict the bipartite worker-firm graph to the largest connected set linked by at least one firm mover (Abowd et al., 2002). Columns (3) and (4) show results for separate connected sets of natives and immigrants. For natives the connected set covers 94.4% and for immigrants 96.1% of individuals in the full sample. The large coverage is made possible by the sample essentially covering the population. The restriction has a bigger effect on the number of firms. The connected set only covers 70.9% and 79.5% of firms for natives and immigrants respectively. This is partly driven by the number of self-employed observations decreasing.

[8] This refers to the year 2019, including the United Kingdom. Actually Switzerland itself is not a member of the European Economic Area but participates in the common market.

Table 1
Summary Statistics

	Full Sample		Connected Set		Dual-Connected Set		
	Natives	Immigrants	Natives	Immigrants	Natives	Immigrants	2000–2004 Arrival Cohort
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Observations:							
Number of person-year obs.	38,599,454	21,540,339	37,000,684	21,027,094	33,262,819	19,696,913	2,867,547
Number of persons	3,567,245	2,647,194	3,367,913	2,544,017	3,222,798	2,452,404	252,255
Number of firms	985,440	580,353	698,989	461,324	222,988	222,988	108,917
Earnings and Firms:							
Mean Earnings	7,423	6,967	7,428	6,976	7,479	7,041	7,379
Mean Log Earnings	8.76	8.67	8.77	8.68	8.78	8.69	8.71
Mean Firm Size	3,009	1,714	3,139	1,756	3,492	1,874	1,810
Gender and Age:							
Male (%)	59.0	63.5	58.4	63.5	57.9	63.1	62.6
Mean Age	42.4	40.9	42.1	40.8	42.1	40.8	38.8
30 or younger (%)	16.6	16.4	17.1	16.6	17.4	16.7	17.2
31 to 49 (%)	53.9	61.8	54.4	62.1	54.1	62.0	71.5
50 or older (%)	29.5	21.8	28.5	21.3	28.5	21.3	11.4
Education:							
Primary (%)	4.5	17.9	4.5	18.1	4.6	18.3	17.8
Secondary (%)	37.6	22.3	38.2	22.5	38.8	23.1	21.4
Tertiary (%)	30.3	22.3	31.0	22.5	32.1	23.2	27.9
Missing (%)	27.6	37.6	26.3	36.9	24.5	35.4	32.9
Origin:							
Balkans and East. Eur. (%)		18.9		18.9		18.5	11.9
Southern Eur. (%)		26.9		26.9		26.6	23.7
North. and West. Eur. (%)		40.0		40.0		41.0	49.1
Other (%)		14.2		14.1		13.9	15.4
Residence Status:							
Foreign Resident (%)		62.7		62.7		62.4	67.3
CBW (%)		19.7		19.9		20.3	24.7
Naturalized Immigrant (%)		3.9		3.9		3.9	5.3
Missing (%)		13.7		13.4		13.4	2.7
Immigration Information:							
Year of Immigration obeserved (%)		99.4		99.4		99.4	100.0
Mean Immigration Age		24.8		24.8		24.7	29.5
Underage at Immigration (%)		20.2		20.0		20.1	3.2
Participating >19 years or 2020 (%)		84.7		85.0		84.0	73.0
Employment Status:							
Employed (%)	93.9	97.7	96.9	98.5	99.7	99.9	99.9
Self-employed (%)	6.1	2.3	3.1	1.5	0.3	0.1	0.1

Notes: Summary statistics for different samples of natives and immigrants. Statistics refer to person-year observations. The connected set restricts the full samples to those workers and firms 'connected' by at least one firm mover for natives and immigrant respectively. The dual connected set restricts the connected sets of natives and immigrants to the firms present in both samples. The immigration year of CBW's is set to the first year of appearance in the full sample. The 2000–2004 arrival cohort contains those immigrants in the dual-connected set immigrating in this time frame.

For being able to draw conclusions between the connected sets of natives and immigrants, it is necessary to define a 'dual-connected set'. The dual connected set restricts the connected sets of natives and immigrants to the firms present in both samples and their respective associated workers. This set does not include firms with either no natives or no immigrants at some point. As a result it covers 90.3% of natives and 92.6% of immigrants. Naturally the dual-connected set contains bigger firms on average, the person-year weighted mean firm size increases by 16% for natives and 9% for immigrants compared to the full sample. Many small firms are dropped, such that firm coverage decreases to 22.6–38.4%. Natives are differentially sorted into firms, their mean firm size is almost twice as large as the immigrant one, $3,492$ vs. $1,874$.

Selectivity in the dual-connected set seems to be rather low, for example natives' (immigrants') mean earnings are roughly 0.7% (1.1%) higher than in the full sample.[9]

Immigrants originate mainly from Northern and Western Europe, with 40%, Southern Europe, with 26.9%, and the Balkans and Eastern Europe, with 18.9% of the person-year observations. 14.2% of the person-year observations are from immigrants born in other countries. Immigrants are overrepresented in the primary education bracket, while natives are overrepresented in the secondary and tertiary schooling levels. In the sample period immigrants are, on average, 1.5 years younger than natives. Looking at residence status, the important role of CBWs in understanding the dynamics of the immigrant-native earnings gap becomes evident. Around 20% of all immigrant person-year observations are based on this group. Relatively many immigrants, 84%, still participate on the Swiss labor market 20 years after their arrival or until 2020. Some of this gap is caused by individuals immigrating already in a high age and not being within working age after 20 years. All these statistics do not differ much over the different samples.

Figure A.6 shows the share of person-year observations by labor market regions separately for natives and immigrants over the sample period for the dual-connected set. Immigrants are especially clustered in the regions of Zurich, Basel, Lausanne and Lugano. Natives are more evenly distributed. Figure A.5 shows the share of immigrant person-year observations by labor market regions in the sample period for the dual-connected set, i.e. the interaction of the two separate distributions. Immigrant workers represent more than 50% of the observations in Geneva and Lugano, underlining their important role in the local labor force. Natives are overrepresented in the regions of the Swiss Plateau and Central Switzerland, including Bern, Fribourg, Biel/Bienne, Aarau-Olten and Lucerne.

Column (8) shows results for the cohort of immigrants within the dual-connected set that arrived between 2000–2004. This is done to isolate assimilation in characteristics from effects caused by newly arriving immigrants. Section 6.2 further elaborates on the composition of this subset.

[9]This could be explained even without a change in the skill-composition of individuals. The observation is consistent with the finding that bigger firms, on average, pay higher wages. See for example Winter-Ebmer and Zweimüller (1999) for evidence in Switzerland.

4 Empirical Strategy

In the this section the econometric framework will be discussed thoroughly. Subsection 4.1 will introduce the AKM-like model used to identify firm premia and describe how it can be used to decompose the variation in log earnings into, among others, components driven by heterogeneous firm wage policies and worker-firm sorting. Thereafter, subsection 4.2 shows how the model can be used to analyze the immigrant-native earnings gap in the cross-section and over time. For the model to be identified, firm premia have to be normalized across groups, this issue is discussed in detail in the end of the section.

4.1 AKM Framework

I decompose log monthly earnings of worker i of group $g \in \{N, M\}$ in year t – $\ln y_{igt}$ – using an AKM-type model, following recent refinements of the approach (e.g. Card, Cardoso, and Kline (2016) and Dostie et al. (2021)):

$$\ln y_{igt} = \alpha_{ig} + \psi^{g}_{j(igt)} + X_{igt}\beta_g + \varepsilon_{igt}, \tag{16}$$

where α_{ig} is a person effect, explaining individual-specific, time-invariant differences in earnings capacity. This could be productivity, but also factors such as negotiation skills. $j(igt)$ is a function identifying the firm of worker i in group g at time t, $\psi^{g}_{j(igt)}$ is a firm effect, explaining firm-specific, time-invariant differences in earnings capacity across all employees with $j(igt) = j$. X_{igt} is a vector of time-varying controls including an age profile, cubic firm tenure[10], residency labor market region fixed effects[11] and time fixed effects. As described by CHK person effects and age linear effects are not separately identified when simultaneously controlling for year fixed effects. This is due to the fact that person effects include cohort effects, as a result age effects can be represented as a linear combination of year and person effects. Therefore the age profile is restricted to be flat at age 50, afterwards I control for a cubic polynomial in age. The age profile is relatively flat at age 50 for natives and immigrants of both genders as can be seen in figure A.1.[12] Consequently person effects represent the earnings capacity at the peak of

[10]Although labor market experience since 1981 can be observed, it only applies to experience on the Swiss labor market, ignoring any experience in other countries. Including it as a covariate leads to heavily positively biased person AKM effects for immigrants (after normalization, see section 4.2).

[11]The allocation of immigrants across regions differs substantially from the native allocation, see figure A.6. Furthermore, tax rates differ considerably across cantons and municipalities, and depend on the place of residence, which could be reflected in realized wages and tax induced intra-national mobility, see for example Schmidheiny and Slotwinski (2018) for evidence on the latter.

[12]The earnings profile of native women experiences a clear dip while the immigrants' profile does not. Part of this can likely be attributed to heterogeneous child penalties. Kleven, Landais, and Leite-Mariante (2023) provide evidence on differences of child penalties for natives and immigrants in the US.

the age profile.
The error term is assumed to have the following form:

$$\varepsilon_{igt} = u_{igj(igt)} + \phi_{j(igt)} + r_{igt}, \tag{17}$$

It captures all remaining determinants of earnings, including person-specific job-match effects $u_{igj(igt)}$, transitory shocks $\phi_{j(igt)}$ affecting all workers with $j(igt) = j$ (e.g. employer demand shocks), and idiosyncratic transitory shocks (including measurement error) r_{igt} affecting the worker (e.g. health shocks)[13].
An important feature of the above model is that firm-specific pay premia $\psi^g_{j(igt)}$ of firm j, although constant within group g, are allowed to vary *between* groups.
Unbiased identification of fixed effect estimates using the above model relies on worker mobility being uncorrelated with time-varying residual components of earnings defined in equation 17. CHK and Card, Cardoso, and Kline (2016) have developed standard diagnostic tests to demonstrate that this *conditional exogenous mobility* assumption holds. A more thorough discussion follows in section 5.1.
Standard AKM-type models assume that firm AKM effects are constant over the sample period. Lachowska et al. (2023) provides evidence showing that firm AKM effects are highly persistent over long time horizons, making a misspecification due to restricting firm effects to be fixed over time unlikely. For this reason, utilization of methods allowing time-varying firm AKM effects à la Lachowska et al. (2023) is not considered in this thesis.

Using the AKM framework model, it is possible to decompose the variance of log monthly earnings:

$$\begin{aligned} var(\ln y_{igt}) = \quad & var(\alpha_{ig}) + \overbrace{var(\psi^g_{j(igt)})}^{\text{firm component}} + var(X_{igt}\beta_g) + var(\varepsilon_{igt}) \\ & + \underbrace{2cov(\alpha_{ig}, \psi^g_{j(igt)})}_{\text{sorting component}} + 2cov(\alpha_{ig}, X_{igt}\beta_g) + 2cov(\psi^g_{j(igt)}, X_{igt}\beta_g). \end{aligned} \tag{18}$$

The above decomposition gives insights on the role of firms in overall earnings inequality. The components $var(\psi^g_{j(igt)})$ and $2cov(\alpha_{ig}, \psi^g_{j(igt)})$ are measures for firm contribution and worker-firm sorting respectively and have been computed in the literature for different time periods and countries.[14]

[13] The identification of the AKM effects is maintained if r_{igt} is allowed to have a unit root as long as it has mean zero within the same individual i, this could reflect for example persistent health shocks, unobserved human capital accumulation (CHK), or even a persistent change in working hours. In this case the zero mean restriction defines the person AKM effect α_{ig} (CHK).

[14] Prior usage includes countries such as France (Abowd et al., 2002), Brazil (Alvarez et al., 2018; Gerard et al., 2021; Engbom and Moser, 2022; Lopes de Melo, 2018), Germany (Card, Heining, and Kline, 2013; Goldschmidt and Schmieder, 2017), Portugal (Card et al., 2018; Card, Cardoso, and Kline, 2016), Austria (Gruetter and Lalive, 2009), Italy (Kline, Saggio, and Sølvsten, 2020; Iranzo, Schivardi,

Estimation errors in the fixed effects, depending on the sparsity of the worker-firm network pose a major problem to the unbiasedness when using sample moments as estimates of the population moments in equation 18. A discussion and an alternative approach can be found in section 5.2.

4.2 Decomposition of the Immigrant-Native Earnings Gap

Decomposition. To analyze the immigrant-native earnings gap I employ a decomposition which has already been applied to analyze firm-driven earnings gaps by gender[15], race[16] and immigrant status[17] using AKM-type models. First let us define $\mathbb{1}_{igt}$ as an indicator function equaling 1 if individual i of group g is employed, i.e. has earnings exceeding the threshold, in year t.

Taking expectations of equation 16 and assuming $\mathbb{E}[\varepsilon_{igt}] = 0$ the mean immigrant-native earnings gap at time t can be represented as:

$$\begin{aligned} \mathbb{E}[\ln y_{iNt}] - \mathbb{E}[\ln y_{iMt}] = \; & \mathbb{E}[\alpha_{iN} \mid \mathbb{1}_{iNt} = 1] - \mathbb{E}[\alpha_{iM} \mid \mathbb{1}_{iMt} = 1] \\ & + \bar{X}_{Nt}\beta_N - \bar{X}_{Mt}\beta_M \\ & + \underbrace{\sum_j \psi_j^N \pi_{Njt} - \sum_j \psi_j^M \pi_{Mjt}}_{\text{firm effect}}, \end{aligned} \tag{19}$$

where $\bar{X}_{gt} = \mathbb{E}[X_{igt} \mid \mathbb{1}_{igt} = 1]$ and $\pi_{gjt} = \frac{\sum_i \mathbb{1}\{j(igt)=j\}}{\sum_i \mathbb{1}_{igt}}$ is the fraction of employees of group g employed by firm j in year t. The first term, the difference in mean person AKM effects of participating individuals, captures changes in the composition of permanent worker skill in the two groups such as selective emigration of low-skilled immigrants (see Dustmann and Görlach (2015)). The second term, the difference between observable time-varying worker characteristics, captures effects such as overrepresentation of immigrants in high-earnings regions or heterogeneous demographic dynamics between the groups.

The third term of equation 19, hereafter *firm effect*, captures the net contribution of firm wage policy differences and can be further analyzed using an Oaxaca (1973)-style decomposition:

and Tosetti, 2008), Israel (Arellano-Bover and San, 2024), Canada (Dostie et al., 2021), Denmark (Bagger and Lentz, 2019) and the US (Song et al., 2019; Sorkin, 2018) for various time periods.

[15] See Card, Cardoso, and Kline (2016); Gallen, Lesner, and Vejlin (2019); Coudin, Maillard, and Tô (2018); Bruns (2019); Sorkin (2017).

[16] See Gerard et al. (2021).

[17] See Damas de Matos (2017); Dostie et al. (2021); Arellano-Bover and San (2024).

$$\sum_j \psi_j^N \pi_{Njt} - \sum_j \psi_j^M \pi_{Mjt} = \underbrace{\sum_j \psi_j^N (\pi_{Njt} - \pi_{Mjt})}_{\text{sorting effect}} + \underbrace{\sum_j (\psi_j^N - \psi_j^M)\pi_{Mjt}}_{\text{pay-setting effect}} . \tag{20}$$

The first term, the *sorting effect*, represents the difference in expected earnings between immigrants and natives attributable to the allocation of immigrants and natives to high-premium firms (measured by the native firm premia). The second term, the *pay-setting effect*, represents the contribution of differential pay setting within firms.

Measuring Assimilation. To further analyse when, how and through which channels assimilation in firm premia takes place I estimate sorting and pay-setting effects' role in assimilation for a specific cohort of immigrants. As mentioned in section 3 the Swiss data are unique in recording the exact date of immigration. Therefore I refrain from measuring assimilation as the change of equation 20 over time like Dostie et al. (2021) and follow a similar approach as Arellano-Bover and San (2024).
Parallel to Arellano-Bover and San (2024), I estimate the following regression:

$$\hat{\psi}^g_{j(igt)} = \mathbb{1}\{g(i) = M\} \times \left[\sum_{e=0}^{20} \beta_e \mathbb{1}\{E_{it} = e\}\right] + X_{it}\theta + \varepsilon_{it}, \tag{21}$$

where $g(i)$ is a function identifying group g of person i. E_{it} is the time since immigration of immigrant i in year t. X_{it} is a vector of controls, including age, year and region fixed effects.[18] The coefficients $\{\beta_e\}_{e=0}^{20}$ are equivalent to the following expectation:

$$\beta_e = \mathbb{E}[\psi^M_{j(igt)} \mid g(i) = M, E_{it} = e, X_{it}] - \mathbb{E}[\psi^N_{j(igt)} \mid g(i) = N, X_{it}]. \tag{22}$$

By adding and subtracting $\mathbb{E}[\psi^N_{j(igt)} \mid g(i) = M, E_{it} = e, X_{it}]$ it is possible to define sorting and pay-setting effects analogously to equation 20:

$$\begin{aligned}\beta_e = &\underbrace{\mathbb{E}[\psi^N_{j(igt)} \mid g(i) = M, E_{it} = e, X_{it}] - \mathbb{E}[\psi^N_{j(igt)} \mid g(i) = N, X_{it}]}_{\text{sorting}} \\ &+ \underbrace{\mathbb{E}[\psi^M_{j(igt)} - \psi^N_{j(igt)} \mid g(i) = M, E_{it} = e, X_{it}]}_{\text{pay-setting}} .\end{aligned} \tag{23}$$

Notice that the components condition on the vector of observed characteristics X_{it}, consistent with equation 21. Thus, in order to estimate equation 23, firm AKM effects are first residualized using $\hat{\theta}$ which is recovered from estimation of 21. Appendix C.1 gives

[18]The equation assumes the same year, age and regional effects for natives and immigrants.

additional information on how the above moments are estimated in the data. When comparing equations 20 and 23 it is evident that these are almost equivalent, the main difference is a reversed sign, i.e. in accordance with the idea of assimilation the latter decomposition measures the gap between immigrants and natives and not vice versa.
Decomposition 23 is able to measure the contribution of firm wage policies towards immigrant assimilation and decompose it into an effect driven by climbing of the firm ladder (sorting) and an effect driven by disparate within-firm wage policies for the two groups (pay-setting) while controlling for time-varying worker characteristics. To isolate changes in sorting and pay-setting effect originating from actual assimilation and changes driven by differential initial allocation and compositional differences of newly arriving immigrants (Borjas, 1985), the above decomposition is applied to a cohort of immigrants arriving between 2000 and 2004, ignoring all later immigrants.
To measure assimilation in all earnings relative to natives I estimate the following regression:

$$\ln y_{igt} = \mathbb{1}\{g(i) = M\} \times \left[\sum_{e=0}^{20} \beta_e \mathbb{1}\{E_{it} = e\} \right] + X_{it}\theta + \varepsilon_{it}, \tag{24}$$

where the parameters, vectors and functions are equivalent to equation 21. Here the $\{\beta_e\}_{e=0}^{20}$ give the expected excess earnings of immigrants relative to natives adjusted for year, age and regional effects. This makes it possible to compute the contribution of firm policies towards earnings assimilation.

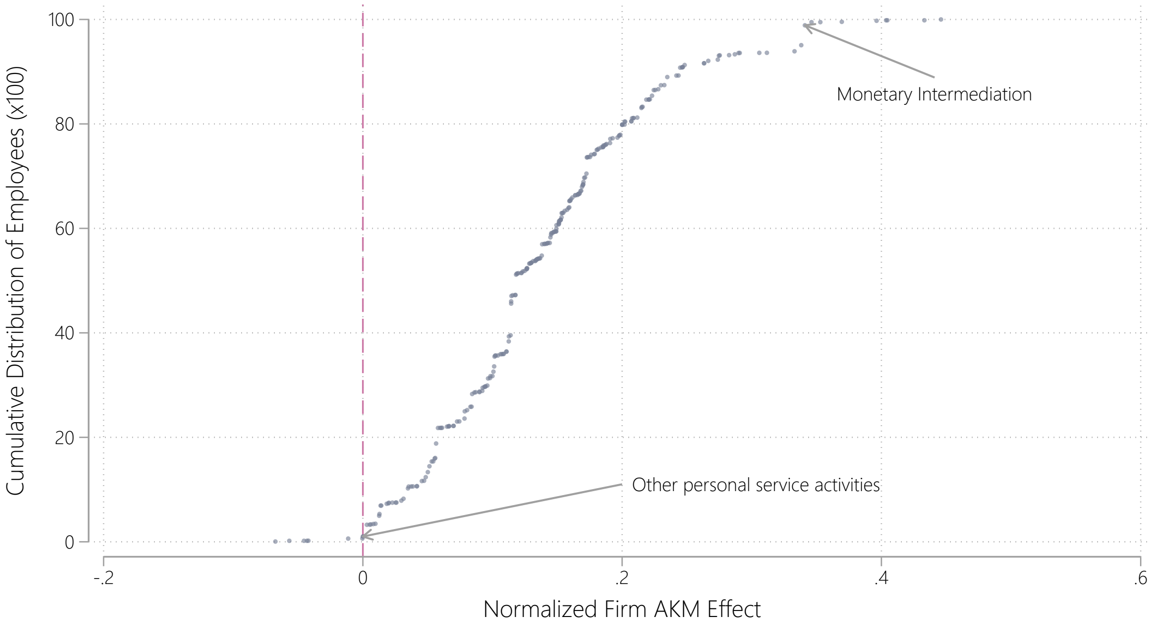
Normalization. The person and firm AKM effects in equation 16 are only identified *within* a connected set linked by workers moving between firms (Abowd et al., 2002), in this case immigrants and natives. Even within the connected set the firm AKM effects identify the firm wage premium relative to a reference firm (Card, Cardoso, and Kline, 2016). Intuitively, subtracting a constant from every person AKM effect and adding the same constant to every firm AKM effect in equation 16 would not change the errors and the fit of the model. To compare firm AKM effects between groups, a normalization across groups is necessary. I follow the literature and normalize the firm AKM effects for both groups g to be zero on average in an industry with low profits. For example, Card, Cardoso, and Kline (2016), Gerard et al. (2021) and Arellano-Bover and San (2024) use the restaurant industry.
The rationale is that the monopsonistic wage-setting model outlined in section 2 predicts that firm premia are proportional to value added per worker. In particular $\psi_j^g = \delta_g R_j$, where R_j is the marginal revenue product of a worker in firm j relative to the reservation wage and δ_g is a group-specific utility parameter influencing the elasticity of labor supply of group g. As shown in the derivation of the model, 'marginally efficient' firms with

$R_j = 1$ will post a wage equal to the marginal productivity of the respective group. Figure A.2 shows the average full-time equivalent productivity, i.e. value added divided by labor input, by 2-digit industry in the sample period. According to the Federal Statistical Office (2023a) the 'Other personal service activities' 2-digit industry, which includes services such as washing of textiles and operating of saunas and tanning studios, has the lowest average productivity if the 'Agriculture, forestry and fishing' industry is omitted which has very low employment, as can be seen in figure A.3, and is likely regionally concentrated. I assume that in this industry, on average, marginal revenue product equals the reservation wage, so workers are paid their marginal productivity. Therefore, firm AKM effects are normalized to be zero for this industry for both, immigrants and natives. Economically this means that all differences in wages in this industry are due to idiosyncratic differences in worker productivity, i.e. person AKM effects or worker characteristics. Interestingly the 'Food and beverage service' industry, which includes restaurants, has a slightly higher, but very similar productivity.

As stated by Dostie et al. (2021), the normalization has important consequences for the interpretation of the immigrant-native earnings gap decomposition in section 6.1. Let the estimated contribution of the firm premia to the wage gap be x ppts. As mentioned above, the normalization assumes that there is, on average, no firm premium for neither group in the 'Other personal service activities' industry. If the assumption does not hold and even firms in low-profit industries pay a premium to natives of $p\%$, this would imply the actual contribution of firm premia to be $x + p$ ppts. This illustrates that the applied normalization gives a conservative, lower bound estimate.
Further, the normalization influences the pay-setting, but *not* the sorting effect estimates from equations 20 and 23. To see this, let us define $\tilde{\psi}_j^N = \psi_j^N + c$. The pay-setting effect in equation 20 becomes $\sum_j (\tilde{\psi}_j^N - \psi_j^M)\pi_{Mjt} = \sum_j (\psi_j^N - \psi_j^M)\pi_{Mjt} + c$, since $\sum_j \pi_{Mjt} = 1$. On the other hand, the sorting effect is unchanged, $\sum_j \tilde{\psi}_j^N(\pi_{Njt} - \pi_{Mjt}) = \sum_j \psi_j^N(\pi_{Njt} - \pi_{Mjt})$, since $\sum_j c(\pi_{Njt} - \pi_{Mjt}) = 0$ for any c (Gerard et al., 2021). This can be shown analogously for equation 23. The normalization *does not* influence the change of the effects, because the intercept of the firm effects is differenced away.

Figure 1 shows the cumulative distribution of normalized firm AKM effects grouped by 3-digit industries. As described, the mean of firm AKM effects for the baseline 'Other personal service activities' industry is approximately zero. Workers in industries like 'Monetary Intermediation', including commercial banks and central banking, experience a firm premium of ≈ 35 log points, i.e. $\approx 42\%$, relative to workers in the baseline industry, conditional on time-invariant individual characteristics and other covariates.

Figure 1
Cumulative Distribution of Normalized Firm AKM Effects by 3-digit Industry



Notes: Cumulative distribution of normalized firm AKM effects by 3-digit industry. For each 3-digit industry the person-year weighted average normalized firm AKM effect (horizontal axis) and the number of person-year observations (vertical axis) is computed. Based on the dual-connected set.

5 Estimation Results

In this section I first present evidence in support of the assumptions of model 16 in subsection 5.1. Afterwards, subsection 5.2 presents estimates of the model fitted to the connected sets of natives and immigrants described in table 1 and results of the variance decomposition of log monthly earnings shown in equation 18. Further, the bias associated with the variance decomposition is discussed and results of alternative, unbiased estimation methods are interpreted. Then, the sorting component of the variance decomposition is further discussed when computing a measure of assortative matching applicable to subgroups of the connected sets in subsection 5.3.

5.1 Exogenous Mobility and Additive Separability

Exogenous Mobility. Assuming there are I_g individuals, J_g firms, the data are observed for T periods, and each respective period t has N_{tg} observations, there are a total of N_g person-year observations in the data for $g \in \{N, M\}$. The AKM framework 16 can be expressed in matrix notation as

$$\ln y_g = D_g \alpha_g + F_g \phi_g + \beta_g X_g + \varepsilon_g, \tag{25}$$

where $\ln y_g$ is the $[N_g \times 1]$ stacked vector of log earnings, D_g is the $[N_g \times I_g]$ design matrix for person fixed effects, F_g is the $[N_g \times J_g]$ design matrix for firm fixed effects and X is the $[N_g \times k]$ design matrix of controls. Restricting the data to the connected set of workers, ensures that matrices D_g and F_g have full rank (Abowd et al., 2002).
As noted by CHK, unbiased estimation of the parameter vector $[\alpha_g^\intercal \quad \phi_g^\intercal \quad \beta_g^\intercal]^\intercal$ using OLS requires the following standard orthogonality conditions to hold:

$$\mathbb{E}[D_g^\intercal \varepsilon_g] = 0; \quad \mathbb{E}[F_g^\intercal \varepsilon_g] = 0; \quad \mathbb{E}[X_g^\intercal \varepsilon_g]. \tag{26}$$

In the case of AKM models, in particular assumption $\mathbb{E}[F_g^\intercal \varepsilon_g] = 0$ is critical. It implies that there is no correlation between residual earnings and allocation of individuals to firms, (Abowd, McKinney, and Schmutte, 2019). This rules out that job mobility is driven by idiosyncratic job-match effects $u_{igj(igt)}$, which are a component of ε_{igt} (Eeckhout and Kircher, 2011). It is important to mention that AKM-type models *allow* for random match effects, whereas they do not allow for the match effects being systematically informative about future job transitions and vice versa.[19]

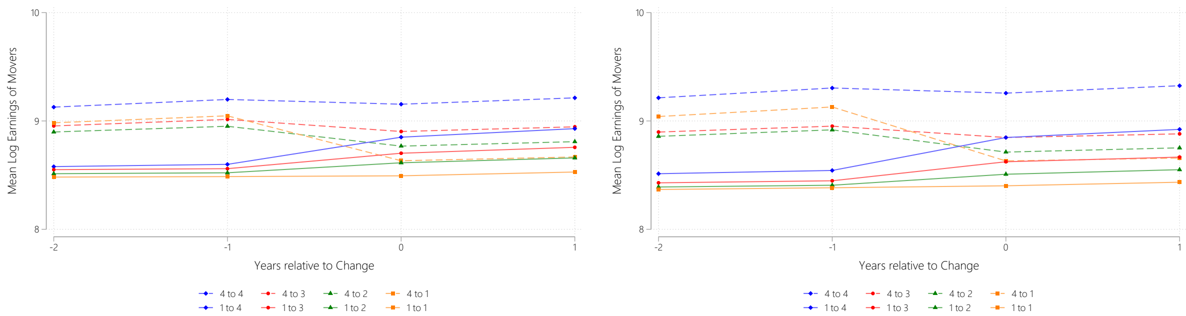
CHK, Card, Cardoso, and Kline (2016) and Card et al. (2018) develop standard diagnostic tests which have been widely applied in the literature.[20] The model implies that individuals moving from high-paying firms to low-paying firms should experience an earnings loss approximately equal to the gain in earnings of individuals moving in the opposite direction. Alternative models where mobility is driven by idiosyncratic job-match effects (e.g. Eeckhout and Kircher (2011)) predict that movers will experience positive gains regardless of the AKM effect of the origin or destination firm.
Figure 2 shows results of the job-move event study proposed by CHK. The figure illustrates the evolution of log monthly earnings around job moves classified by origin and destination firm AKM effect quartiles for both connected sets. For clarity, only the evolution of wages for workers leaving from first and fourth quartile firms are plotted. For both, immigrants and natives, movers staying within their firm AKM effect quartile have a flat earnings profile.

[19] Orthogonality condition $\mathbb{E}[D_g^\intercal \varepsilon_g] = 0$ is less thoroughly discussed in the literature. It states that there is no correlation between residual earnings and an individual's decision to actively participate in the labor market (Abowd, McKinney, and Schmutte, 2019). If, for example, individuals only decide to work when match effects $u_{igj(igt)}$ or firm-wide earnings shocks $\phi_{j(igt)}$ are high, this could bias the estimates. Although it could be argued that this is mainly a concern for estimates of person AKM effects and these are less important in the following analysis. For example, CHK impose assumptions on the structure of the model's residuals, which rule out a violation of the condition.

[20] See e.g. CHK for Germany, Card, Cardoso, and Kline (2016) and Card et al. (2018) for Portugal, Macis and Schivardi (2016) for Italy, Gerard et al. (2021) for Brazil, Song et al. (2019) for the United States and Arellano-Bover and San (2024) for Israel.

Figure 2
Job Move Event Study

(a) Natives **(b)** Immigrants



Notes: The figure shows earnings trends of job movers around job changes from the top and bottom quartiles of co-worker firm AKM effects (person-year weighted quartiles) at origin moving to destination establishments in any of the other quartile groups. Movers are defined as workers who separated from the origin firm in $t = -1$ or during $t = 0$, joined the destination firm in $t = 0$, and were employed at the origin and destination firm for 2+ consecutive years. Based on the largest connected set for each origin group.

The absence of a general mobility premium for these workers suggests that job mobility is not driven by job-match effects $u_{igj(igt)}$ or employer-wide shocks $\phi_{j(igt)}$. Movers climbing the firm ladder systematically experience an increase in earnings, whereas movers descending on the firm ladder systematically have lower earnings after the move. Furthermore, gains of movers going from the first to the fourth quartile are comparable to the losses of movers going in the opposite direction.
This evidence supports the additive nature of the model. There are no obvious trends before or after the job move, and no systematic 'dip' in the period of transition as for example Dostie et al. (2021) and Song et al. (2019) find them using annual earnings data. Overall the patterns are broadly in line with the prediction of the model.

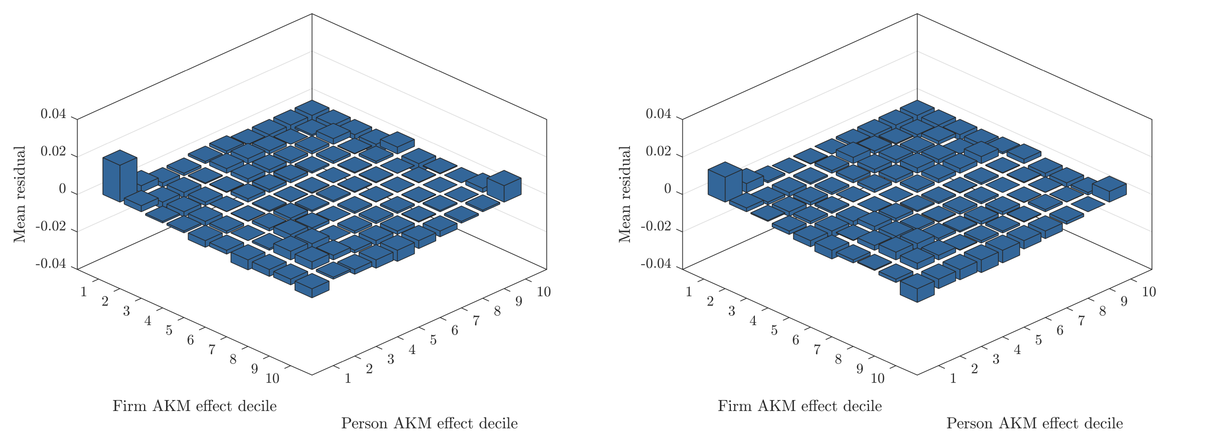
Job-Match Effects. Although the above mentioned event study shows no signs of a significant influence of job-match effects, this can be further investigated. To quantify the importance of job-match effects I follow CHK and fit a job-match effects model to the data. It includes a separate fixed effect $\mu_{igj(igt)}$ for every distinct worker-firm match. To get an estimate of the variance of the job-match component that is included in the AKM error term 17, CHK use $\widehat{var}(u_{igj(igt)}) = MSE_{AKM_g} - MSE_{jobmatch_g}$ as an approximation. Panel A and C of table 2 show the number of fixed effects, adjusted R^2, root mean squared error and the standard deviations of the components for the AKM and job-match effects model respectively. As expected the adjusted R^2 of the job-match effect is higher than in the AKM model, 0.835 and 0.867 vs. 0.883 and 0.899, due to it having additional fixed effects, although the difference is quite low with 3–5% of additional variance explained. The standard deviation of the job-match effect component, calculated as shown above, for natives (immigrants) is 0.109 (0.092), which accounts for 4.8% (3.2%) of the total variance in log earnings. Although this does not show that job-match effects are independent of worker mobility per se, they do not seem to play an important role in wage determination in the sample. This, together with the job-move event study results shows that job-match effects $u_{igj(igt)}$ are unlikely to be a major driver of bias and $\mathbb{E}[F_g^\intercal \varepsilon_g] = 0$ is likely to hold.

Log-Additivity. The AKM approach assumes that earnings are log additive in person and firm fixed effects. Bonhomme, Lamadon, and Manresa (2019) show that earnings in worker-firm panels are approximately log-additive in person and firm fixed using a more general kind of model.[21] Further, it is possible that specific combinations of workers and firms experience especially high job-match effects.

To test the fit of the model, CHK propose analyzing the estimated residuals. Figure 3 shows mean residuals by person × firm AKM effect decile cells. A high mean residual indicates that the log-additive nature of the model systematically underestimates log earnings for the respective cell and vice versa. All residuals are weak in magnitude. Except for the lowest cell there are no obvious outliers.

Figure 3
AKM Residuals by Person and Firm AKM Effect Deciles

(a) Natives **(b)** Immigrants



Notes: Mean AKM model residuals by person-year weighted firm and person AKM effect deciles. Based on the largest connected set for each origin group.

A possible explanation for the discrepancy could be minimum wages, though these have been introduced in only a few cantons and approaching the end of the sample period,

[21]The AKM framework 16 can be seen as a special case of the static model proposed by Bonhomme, Lamadon, and Manresa (2019).

beginning with the canton of Neuchâtel in 2017.[22] The more likely explanation is the effect of collective labor agreements negotiated between employers and employees. Figure B.1 shows the estimated residuals for the sample consisting of only males. Here, as expected, model fit is improved. This could be due to a lower impact of labor supply effects on estimated firm and person AKM effects or collective labor agreements being less binding for men than women. Low and non-systematic residuals, such as in figures 3 and B.1, are typically interpreted as supportive evidence for the log-additivity assumption.

[22]For evidence on the effects of the implementation of a supreme court-ordered, unexpected minimum wage policy in Neuchâtel, see Berger and Lanz (2020)

Table 2
AKM Decomposition

	Samples	
	Natives (1)	Immigrants (2)
Panel A: Largest Connected Set		
Mean of log earnings	8.770	8.678
Standard deviation of log earnings	0.499	0.515
Number of movers	1,955,123	1,351,412
Model Estimates		
Std. dev. of person effects (across person-yr obs.)	0.402	0.400
Std. dev. of firm effects (across person-yr obs.)	0.170	0.176
Std. dev. of covariates (across person-yr obs.)	0.095	0.095
Correlation of person/firm effects	0.055	0.208
Adjusted R-squared	0.835	0.867
RMSE	0.203	0.188
Explained log earnings variance		
Person effect	64.9%	60.2%
Firm effect	11.6%	11.7%
Covariance of person and firm effects	3.0%	11.1%
Covariates and associated covariances	5.8%	5.6%
Residual	14.7%	11.4%
Panel B: Leave-One-Out Set		
Coverage of the connected set	96.2%	96.6%
AKM		
Correlation of person/firm effects	0.128	0.293
Person effect	61.5%	59.6%
Firm effect	8.2%	9.4%
Covariance of person and firm effects	5.9%	13.9%
Covariates and associated covariances	6.0%	5.7%
Residual	14.9%	11.4%
KSS		
Correlation of person/firm effects	0.161	0.334
Person effect	62.8%	57.2%
Firm effect	7.4%	8.7%
Covariance of person and firm effects	6.9%	14.9%
Covariates and associated covariances	3.9%	4.3%
Residual	19.0%	14.9%
Panel C: Job-Match Effect Model		
Number of job-match effects	7,626,043	5,375,976
RMSE of job-match effect model	0.171	0.164
Adjusted R-squared of job-match effect model	0.883	0.899
Standard deviation of job-match effect	0.109	0.092

Notes: Panel A shows results of decomposition 18 for the largest connected set for every origin group. Panel B shows the same results for the leave-one-out-set, i.e. the largest connected set with at least two movers in every firm. Additionally Panel B shows bias-corrected estimates using the method outlined in KSS. The job-match effect model controls for a fixed effect for every person-firm match instead of person and firm fixed effects.

5.2 Firm-Driven Wage Inequality

Variance Decomposition. Panel A of table 2 gives results of decomposition 18, where sample equivalents are used to estimate population moments, the so-called 'plug-in' estimator.

Time-invariant worker characteristics explain 64.9% of the total variance in log monthly earnings for natives and 60.2% for immigrants. Variation in firm wage policies explain roughly 11.6% of earnings variance for both, natives and immigrants. Sorting of high-wage workers to high-wage firms (and vice versa) explains 11.1% of earnings variation, this is much more than the 3.0% for the native sample. This is reflected by the correlation between person and firm AKM effects, which is 0.055 for immigrants vs. 0.208 for natives. Covariates explain only a small share for both groups. The residual component explains 3.3 ppts. more for the natives than for immigrants. Part of this difference could be driven by the 1.6 ppts. higher job-match effects for natives explained in section 5.1.

The Swiss estimates are similar to the estimates for Israel between 1991–2019 by Arellano-Bover and San (2024) who also find that sorting is higher for immigrants than natives (5.8% vs. 12.1% for males). Although the magnitude of the shares differs, parallel to Swiss estimates, they find that to the contribution of the firm AKM effect variation is similar for both groups, 18.5% vs. 20.5% for males, while person effects are more important for natives than immigrants, 56.9% vs. 37.7%. Dostie et al. (2021) estimate relatively similar shares for both groups for Canada between 2005–2013. It is worthwhile to mention that the Swiss data include a much higher share of immigrant observations compared to Dostie et al. (2021) (35.8% vs. 18.0%) and are in general much bigger than the sample of Arellano-Bover and San (2024) (58m vs. 14m).

The estimates are also in line with non-group specific decompositions like CHK who find a sorting share of 2.3–16.4% for West German men between 1985 and 2009 or Card, Cardoso, and Kline (2016) who find 11.4% for male workers in Portugal between 2002 and 2009.

Bias Correction. Estimates of the explained share of log variance should be interpreted cautiously, as initially noted by Abowd et al. (2004) and later developed by, among others, Andrews et al. (2008), Bonhomme, Lamadon, and Manresa (2019) and KSS. Although estimates of ψ_j^g and α_{ig} are unbiased when assumptions 26 hold, they still contain random estimation errors, i.e., for firm AKM effects, $\hat{\psi}_j^g = \psi_j^g + \xi_j^g$, where ξ_j^g follows a normal distribution with $\mathbb{E}[\xi_j^g] = 0$ and $\mathbb{E}[(\xi_j^g)^2] > 0$. If one uses a simple 'plug-in' estimator for the variances, that is, use $var(\hat{\psi}_j^g)$ and $var(\hat{\alpha}_{ig})$ as estimates for the variances of the true AKM effects like in Panel A of table 2, this leads to an upward bias of the respective variances. Intuitively this is because when computing the second moment of the estimate $(\hat{\psi}_j^g)^2$, one not only squares the true parameter $(\psi_j^g)^2$, but also the random noise $(\xi_j^g)^2$ (Lachowska et al., 2023).[23] Inflated variances attenuate the estimated covariance $cov(\hat{\psi}_j^g, \hat{\alpha}_{ig})$. Furthermore, the effects enter the estimation equation 16 additively, thus estimation errors in person and firm AKM effects are negatively correlated. For example, if a firm AKM effect $\psi_{j(igt)}^g$ is overestimated, the corresponding person AKM effect α_{ig} will be underestimated and vice versa (Andrews et al., 2012). This induces an additional negative bias onto the sorting component.

The above mentioned bias can be a major problem for 'thinner', less densely connected worker-firm graphs, as shown by Bonhomme et al. (2023) with data from Austria, Italy, Norway, Sweden and the United States. This can be intuitively explained: if a sub-graph is connected to a larger part of the graph only by one moving worker, all the AKM effects in the sub-graph are identified by this worker. An estimation error in the person and firm AKM effects of this worker will influence all the estimates in the sub-graph. Therefore, sparsely connected bipartite worker-firm graphs are more likely to experience the bias.

There are several informal strategies to reduce the bias, for example by restricting the sample to big firms, which are more likely to have many movers (Song et al., 2019; Bassier, Dube, and Naidu, 2022; Sorkin, 2018). Unfortunately, this can induce sample selection bias, especially because in this case it is possible that immigrants and natives sort heterogeneously to small and big firms.

KSS develop an alternative approach, although to use it the sample must be restricted to a set, which retains connected even when any one worker is omitted, hereafter called the leave-one-out-set.[24]

[23]It is important to mention that this bias only applies to estimates of second moments. The decomposition of the immigrant-native gap using the framework outlined in section 4.2 does not suffer from this bias, as it is based on differences in expectations and not quadratic terms. By conditional exogenous mobility $\mathbb{E}[\hat{\psi}_j^g] = \psi_j^g$ holds.

[24]KSS demonstrate that the 'plug-in' solution gives estimates which are biased involving a linear combination of the unknown variances $\{\sigma_i^2\}_{i=1}^n$. They develop a computationally feasible method to effectively obtain estimates of the unknown variances, $\{\hat{\sigma}_i^2\}_{i=1}^n$, in large data sets using a leave-one-out approach. These can be used to compute unbiased second moments of person and firm AKM effects. The idea is similar to heteroskedasticity-robust standard errors à la White (1980). There exist two more popular approaches to mitigate limited mobility bias which are used frequently. First, Andrews et al. (2008) propose a similar method to KSS, which assumes homoskedastic errors. Second, Bonhomme, Lamadon, and Manresa (2019) propose a two-step group fixed effect approach (BLM model), where in a first step $k \in \mathbb{N}^+$ groups of firms are constructed according to their empirical wage distribution using k-means classification. In a second step the an AKM-type estimation is conducted using the clustered group identifiers instead of firm identifiers. Obviously, this leads to a densely connected set, because there are more movers between the few clusters than thousands of firms. Bonhomme et al. (2023) build on this estimator and use a correlated random effects (CRE) model in the second stage. Bonhomme et al. (2023) compare the different bias correction methods and conclude that all perform reasonably well and there are no large differences. Estimates of the contribution of firms towards earnings inequality using the CRE and a static version of the BLM model, both with $k = 10$, can be found in table A.2. The estimates using CRE and BLM are very similar and show the same pattern as the KSS estimates, but attribute approximately 5 ppts. more of the variance in log earnings towards the sorting component.

Panel B of table 2 presents results of the KSS decomposition applied to Swiss data. For comparison 'plug-in' estimates for the leave-one-out set are given.[25]
The leave-one-out sets cover more than 96% of the connected set for both groups, estimates should therefore be representative. Although firm AKM effects decrease in importance and sorting increases in importance, the difference between 'plug-in' estimates of the connected and leave-one-out set and the KSS-corrected estimates is relatively small in magnitude. This is consistent with Lachowska et al. (2023), who argue that the KSS correction is of major importance for short panels, whereas long panels take advantage of a higher mover to firm ratio, which mitigates sampling errors and therefore makes 'plug-in' estimates more aligned with KSS estimates. Nevertheless, the preferred estimate of firm influence and worker-firm sorting is the KSS approach. According to it, firms and sorting components taken together are responsible for 14.3% of the variation in log monthly earnings for natives and 23.6% for immigrants.
Bias correction is computationally expensive and not straightforward to implement. Other contributions to the literature of decomposing the immigrant-native gap using AKM-type models, namely Dostie et al. (2021) and Arellano-Bover and San (2024), do not show results, prohibiting a comparison.

Pooled Estimates. Comparison of estimates of equation 18 across countries can give an idea about differences in the importance of firms and the labor market in general. Most other studies do not expand the AKM framework to include group-specific firm AKM effects, which makes comparison across studies problematic. For this reason table A.4 gives 'plug-in' and KSS estimates of the variance decomposition for a mixed set of workers, without group-specific firm AKM effects. The KSS estimates, 10.2% sorting share and 7.6% firm component share, are in line with bias-corrected estimates by Bonhomme et al. (2023) for Austria, Italy, Norway, Sweden and the United States for different periods, where sorting explains 5.0–13.0% and firm AKM effect explains 5.8–15.7% of the variation in log earnings.[26]

Working Hours. Bonhomme et al. (2023) compute the decomposition for both, hourly and annual wages, using Norwegian data. They find that bias-corrected estimates are quite similar across both income measures. For the Swiss case, where monthly earnings are observed, it can be expected that a decomposition of hourly wages would give similar results. As explained earlier, labor supply effects are less likely to play a major role in the

[25] The computation of KSS and CRE models was performed in Python using the *pytwoway* package associated with Bonhomme et al. (2023). The package can be accessed at `https://github.com/tlamadon/pytwoway`.

[26] Bonhomme et al. (2023) exclude public sector employees from the analysis.

person and firm AKM effects when using a sample consisting of only the male population. Table B.1 shows results of the 'plug-in' decomposition, as well as the KSS estimates for the set of male individuals. Firm and sorting components are virtually unchanged.

Regions. The baseline results control for residency in one of 16 labor market regions. Table A.1 shows estimates of regional fixed effects relative to Geneva, figure A.7 shows these on a map.[27] As can be seen, average earnings differ widely over regions, especially so for natives. Earnings for natives are highest in the Geneva and Zurich regions, while natives in the Bellinzona region suffer from a 5.7% average penalty on their earnings relative to Geneva. Interestingly, immigrants' region fixed effects have a much more narrow distribution. When excluding the Zurich region, which has a premium of 3.3% relative to Geneva, all remaining estimates are in a range of 1.7 ppts.
Alternatively, spatial allocation could be considered a channel of labor market, cultural or institutional discrimination. This would be the case if, for example, immigrants are not able to move to economically prosperous areas because they cannot find a job, integration is hindered, or immigrants are not granted the right to move to these places. For this reason, table A.3 reports the decomposition results without including regional controls, here regional effects are partly absorbed into person and firm AKM effects based on the residence of individuals and spatial worker composition of firms over time. The results are robust to this change.

5.3 Assortative Matching

To measure sorting for different subgroups, in theory it is possible to construct connected sets for each subgroup (e.g. age, origin) and estimate the KSS correction. However, this approach would lead to even more sparsely connected sets and, as shown by Bonhomme et al. (2023), the KSS approach looses its ability to mitigate limited mobility bias if the share of movers kept is sufficiently low. Therefore, I follow Dostie et al. (2021) and regress the person AKM effect of person i of group g in year t onto the person's group-specific firm AKM effect in year t:

$$\hat{\alpha}_{igt} = \gamma_0 + \gamma_1 \hat{\psi}^g_{j(igt)} + \eta_{igt}. \tag{27}$$

As mentioned above, sampling errors in person and firm AKM effects are negatively correlated, so OLS estimates are expected to be negatively biased. Additionally, measurement errors in $\hat{\psi}^g_{j(igt)}$ lead to attenuation bias in $\hat{\gamma}_1$. To overcome this, the other group's firm

[27] Estimates of the other covariate coefficients, namely age and tenure, can be found in figure A.8.

effect is used as an instrument in a 2SLS estimation.[28]

Table 3 presents OLS and 2SLS estimation results of γ_1 for different subgroups of natives and immigrants in the dual-connected set. Standard errors are clustered at the person level. As expected, OLS estimates are negatively biased. Overall, firm AKM effects and person AKM effects are highly correlated, which indicates a high degree of assortative matching. Parallel to the findings of Dostie et al. (2021) for Canada, Immigrants experience higher assortative matching than natives in Switzerland. A firm offering a 1% higher firm premium for natives (immigrants) will, on average, have a native (immigrant) workforce with a 0.852% (1.208%) higher time-invariant earnings capacity.

Table 3
Assortative Matching

	OLS		2SLS	
	Natives (1)	Immigrants (2)	Natives (3)	Immigrants (4)
All	0.485***	0.751***	0.852***	1.208***
By Age:				
30 or younger	0.0823***	0.311***	0.394***	0.604***
31 to 49	0.534***	0.840***	0.863***	1.279***
50 or older	0.616***	0.812***	1.092***	1.497***

Notes: Standard errors are clustered at the individual level. For reasons of clarity they are not reported. *** indicate significance at $p < 0.001$. Based on the dual-connected set.

The coefficient increases monotonically, but with diminishing 'marginal effect' with age for both groups. Although this observation could be driven by cohort effects, it is consistent with the idea that 'good' workers need some time to climb the firm ladder through signalling and 'good' firms need information on job history to screen for 'good' workers.[29] These results mirror the higher sorting shares found for immigrants in table 2. Further research on the level of assortative matching over time and the influence of cohort effects follows in section 7.

[28]See Jochmans and Weidner (2019) for a discussion on inference using fixed effects estimated from network data as regressors.

[29]Sorting of workers to high-premium firms over time or based on time-invariant worker characteristics does not violate the orthogonality assumptions 26 as I condition on person, firm and age effects in equation 16.

6 Immigrant-Native Earnings Gap

This part begins by section 6.1 analyzing the role of firms in the immigrant-native earnings gap in terms of sorting and pay-setting effects outlined in equation 20. Next, changes of sorting and pay-setting effects over time are analyzed in section 6.2 to measure their contribution towards immigrant assimilation. This is done using the framework shown in equation 23.

6.1 Earnings Gap in the Cross-section

Table 4 presents results of decomposition 20 estimated on the dual-connected set in the cross-section, i.e. assuming t spans the time period 2002–2020. Column (1) reports the mean difference in log monthly earnings between natives and immigrants across all periods, Columns (2) and (3) decompose that difference into contributions of person AKM effects and covariates.

Table 4
AKM Earnings Gap Decomposition

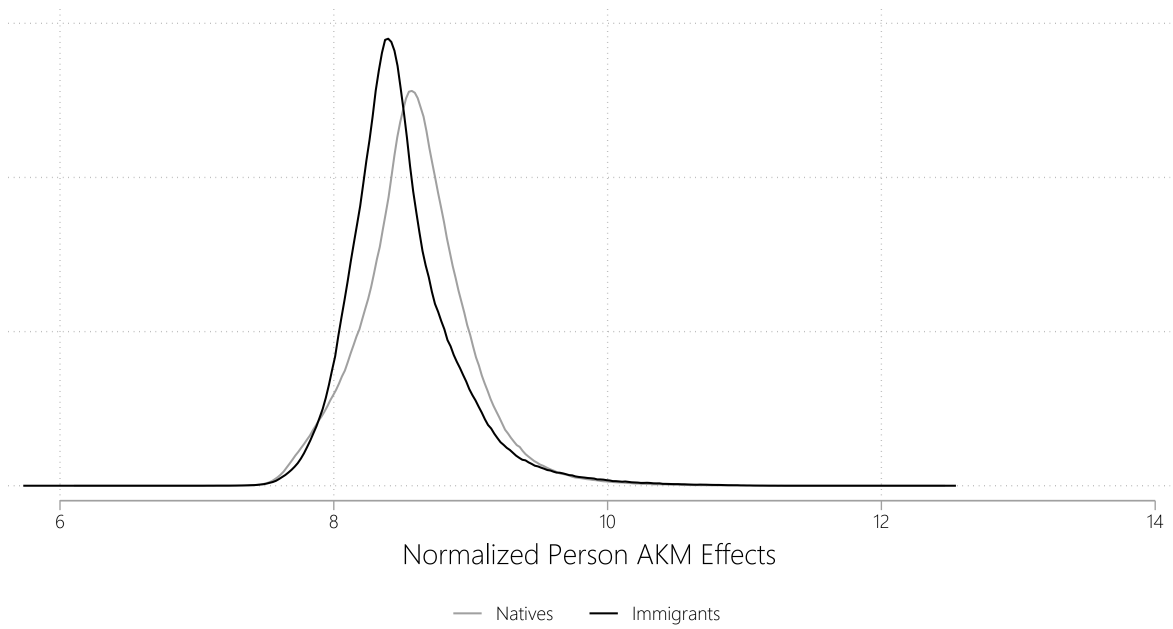
		Transferable Skills		Firms				
	Differece in Log Earnings (1)	Differece in Person Effect (2)	Differece in Log Covariates (3)	Swiss Firm Effect (4)	Immigrant Firm Effect (5)	Firm Effect (6)	Sorting (7)	Pay-setting (8)
All	0.095	0.094 98.9%	-0.021 -22.1%	0.143	0.122	0.021 22.1%	0.005 5.3%	0.016 16.8%
By Gender:								
Male	0.176	0.174 98.9%	-0.020 -11.4%	0.154	0.131	0.023 13.1%	0.013 7.4%	0.009 5.1%
Female	0.014	0.012 85.7%	-0.021 -150.0%	0.127	0.106	0.021 150.0%	-0.007 -50.0%	0.029 207.1%
By Age:								
30 or younger	0.085	0.065 76.5%	-0.013 -15.3%	0.133	0.101	0.032 37.6%	0.013 15.3%	0.019 22.4%
31 to 49	0.100	0.102 102.0%	-0.023 -23.0%	0.148	0.127	0.021 21.0%	0.005 5.0%	0.016 16.0%
50 or older	0.079	0.095 120.3%	-0.032 -40.5%	0.138	0.123	0.015 19.0%	0.001 1.3%	0.014 17.7%
By Education:								
Tertiary Education	-0.040	-0.009 22.5%	-0.024 60.0%	0.161	0.168	-0.007 17.5%	-0.032 80.0%	0.025 -62.5%
No Tertiary Education	0.119	0.113 95.0%	-0.027 -22.7%	0.139	0.106	0.033 27.7%	0.018 15.1%	0.015 12.6%

Notes: Decomposition of immigrant-native earnings gap using equations 19 and 20. Based on the dual-connected set.

Columns (4) and (5) show mean firm AKM effects for natives and immigrants, while column (6) gives the difference between those, i.e. the firm effect. Columns (7) and (8) further decompose the firm effect into contributions of sorting and pay-setting effects. Negative numbers indicate that the respective component alters the earnings gap in favor of immigrants. Percentages under the estimates give the shares of the earnings gap that can be explained by the respective component. The sign of the percentages specifies if the component widens the earnings gap (positive) or attenuates it (negative). Additionally the sample is split by gender, age and education subgroups, the rows compare immigrants to natives of the *same* gender, age or education level.

Aggregate. When comparing all immigrants to all natives in the cross-section, the overall earnings gap is 9.5 log points, differences in person AKM effects explain most of the discrepancy (98.9%). This can be visualized when comparing estimated densities of person AKM effects of the groups in figure 4. Clearly the natives' distribution is shifted to the right. Interestingly the natives' density has a fatter left tail which even overlaps with the immigrants' distribution. This can likely be attributed to part-time working female workers. Estimating the same densities with the connected set based on only male workers gives the densities depicted in figure B.2, where the left tail is not evident and the shift is more prominent. Time-variant worker characteristics, spatial allocation and heterogeneous macroeconomic shocks mitigate the earnings gap by 22.1%.

Figure 4
Density of Person AKM Effects



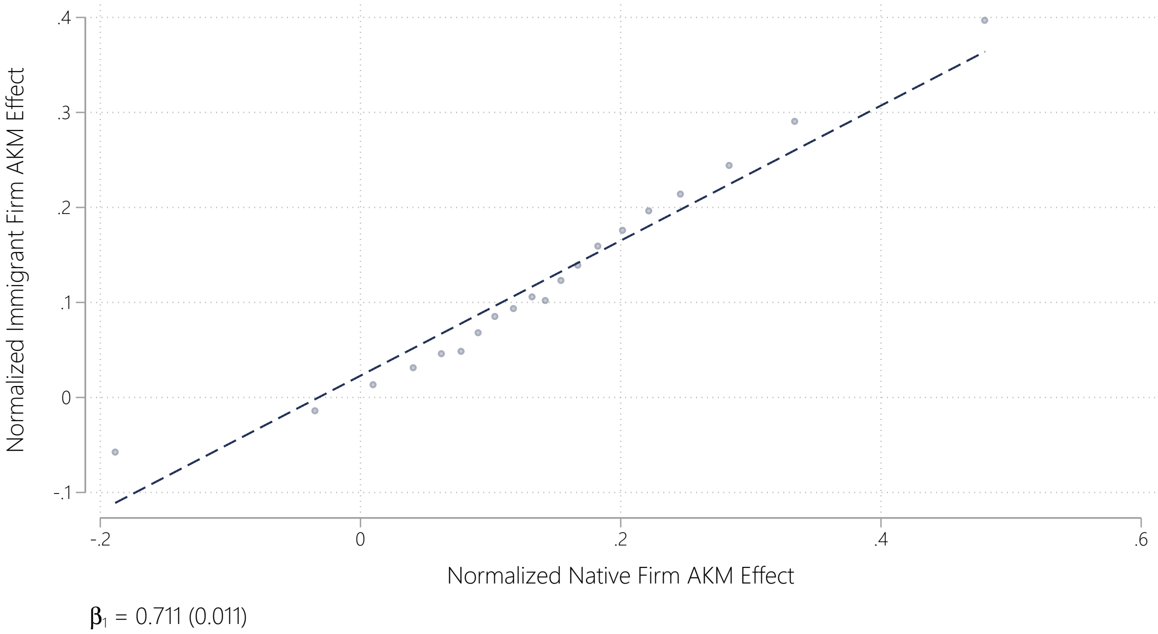
Notes: Estimated person-year weighted density of normalized person AKM effects by origin using epanechnikov kernel and optimal bandwidth. Based on the dual-connected set.

Overall, person AKM effects and time-varying characteristics, which are transferable when switching employers account for 76.8% of the earnings gap.

The mean firm premium relative to the baseline industry for natives is 14.3%, while it is 12.2% for immigrants. The difference in mean firm premia, the *firm effect*, amounts to 2.1 ppts. and accounts for 22.1% of the overall earnings gap. Decomposing the firm effect using equation 20 shows that differential sorting of immigrants and natives to high-paying firms accounts for 5.3% of the earnings gap, while the pay-setting effect accounts for 16.8%. Almost all of the difference in firm AKM effects is attributable to the pay-setting channel, where the same firm pays different premia to natives and immigrants. Computing the same composition on the connected set of males gives even stronger estimates of the sorting and pay-setting effect of 5.8% and 29.5% respectively (see table B.2).

Figure 5 shows a binned scatter plot of $\hat{\psi}_j^M$ vs. $\hat{\psi}_j^N$. The OLS estimate is 0.711. Arellano-Bover and San (2024) estimate the same coefficient and get very similar results with 0.71 for males and 0.70 for females.

Figure 5
Immigrant vs Native Firm AKM Effects



Notes: Binned scatter plot of normalized firm AKM effects of immigrants vs natives. Fitted line estimated using OLS. Estimated standard error given in parenthesis. Standard errors clustered at firm level. Estimation conducted on the person-year level. Based on the dual-connected set.

As I computed KSS estimates of the true variance in firm AKM effects (for the leave-one-out set), it is possible to approximately correct for attenuation bias caused by measurement error in $\hat{\psi}_j^N$. Using this back-of-the-envelope calculation one gets $\hat{\beta}_{corrected} =$

$\hat{\beta}_{OLS}\frac{var(\hat{\psi}_j^N)}{var(\hat{\psi}_{jKSS}^N)} \approx 0.711 * \frac{0.01977}{0.01793} = 0.784$.[30] An immigrant switching to a better-paying firm will only benefit of roughly 78% of the earnings increase a native would experience doing the same transition. Although this is only a rough estimate, it is broadly in line with the estimates of the pay-setting effect in table 4 of 16.8%. A coefficient below 1 means that the firm premia of immigrants are compressed relative to the native ones, i.e. natives have larger firm ladder steps. This is in line with lower firm-specific labor supply elasticities of immigrants in the monopsonistic wage setting model outlined in section 2, where the above estimate corresponds to the relation (δ_M/δ_N), i.e. the quotient of relative valuations of excess wage to non-pecuniary amenities between immigrants and natives (Gerard et al., 2021).

As mentioned in section 4.2, the sorting effect is invariant to the normalization of AKM effects across connected sets, while the firm effect and the pay-setting effect react 1:1 to a change of the baseline firm AKM effect. I regress log hourly earnings from the ESS on an immigrant dummy, controlling for gender, region, age interacted with gender, time, education, 2-digit occupation and an index for managerial responsibility using sample weights. The estimated earnings gap is approximately -6.7%. If the immigrant-natives earnings gap in the baseline industry was due to differential pay-setting rather than differences in worker productivity this would *increase* the pay-setting effect by 6.7 ppts., therefore the presented estimate provides a conservative lower bound.

Gender. The influence of firm-specific wage policies on the gender wage gap has been a important topic of recent literature (see e.g. Card, Cardoso, and Kline (2016) or Sorkin (2017)). Rows two and three of table 4 show estimates separately for males and females. Differences in log earnings between natives and immigrants are much more pronounced for men than women, although this could be driven by higher working hours of immigrant women relative to native ones. Despite this, the immigrant-native gap in firm AKM effects is similar across gender (2.1–2.3 ppts.). For women, this is almost entirely driven by unequal wage policies, while for men, sorting to high-premium firms explains more than half of the gap.

For both, immigrants and natives, the average firm premia are 2.5–2.7 ppts. higher for men than for women. This explains 6.1% of the 41 log point[31] and 10.8% of the 25 log point unadjusted gender wage gap for natives and immigrants respectively. These estimates are lower than the numbers obtained by Card, Cardoso, and Kline (2016) who find

[30] The estimate for $var(\hat{\psi}_j^N)$ is estimated for the leave-one-out set. The estimate for $var(\hat{\psi}_{jKSS}^N)$ is not reported in table 2. It would be more appropriate to compare the firm AKM effect variance on the connected set with a corrected version based on the connected set itself. The difference between these is expected to be bigger, increasing the correction factor and $\hat{\beta}_{corrected}$. Therefore the above correction still likely underestimates the true β, nevertheless it gives a lower bound.

[31] The unadjusted gender wage gap is not shown in the table.

estimates in the range of approximately 20% for Portuguese data.
Note that there are two caveats when comparing the above analysis to Card, Cardoso, and Kline (2016): The used earnings measure does not control for differences in working hours between males and females, presumably overstating gender wage gaps and understating the relative contribution of firm wage policies (Dostie et al., 2021). Furthermore I do not estimate separate firm AKM effects by gender, consequently gender-specific wage policies are prohibited, which have been identified as an additional driver of the wage gap by Card, Cardoso, and Kline (2016). Dostie et al. (2021), using the same approach as me, estimate more similar contributions of 12 and 17% for immigrants and natives respectively for Canadian data.

Age. It is possible to compute the decomposition for different subgroups of the dual-connected set. The contribution of the firm effect is substantial, for workers age '30 or younger', it explains 37.6% of the total earnings gap. With age the contribution decreases to 19% for ages '50 or older'. Interestingly, the majority of this decrease can be attributed to a decrease in the sorting component, which decreases from 15.3% to 1.3%. Furthermore, almost all of the decrease takes place between the age brackets '30 or younger' and '31 to 49'. The pay-setting effect, although decreasing slightly, stays relatively consistent at 16–22.4%. It should be pointed out that the analysis does not control for cohort effects, the change in the importance of the effects could thus be attributed to a differences in immigrant composition over time. A thorough discussion of cohort effects follows in section 6.2.

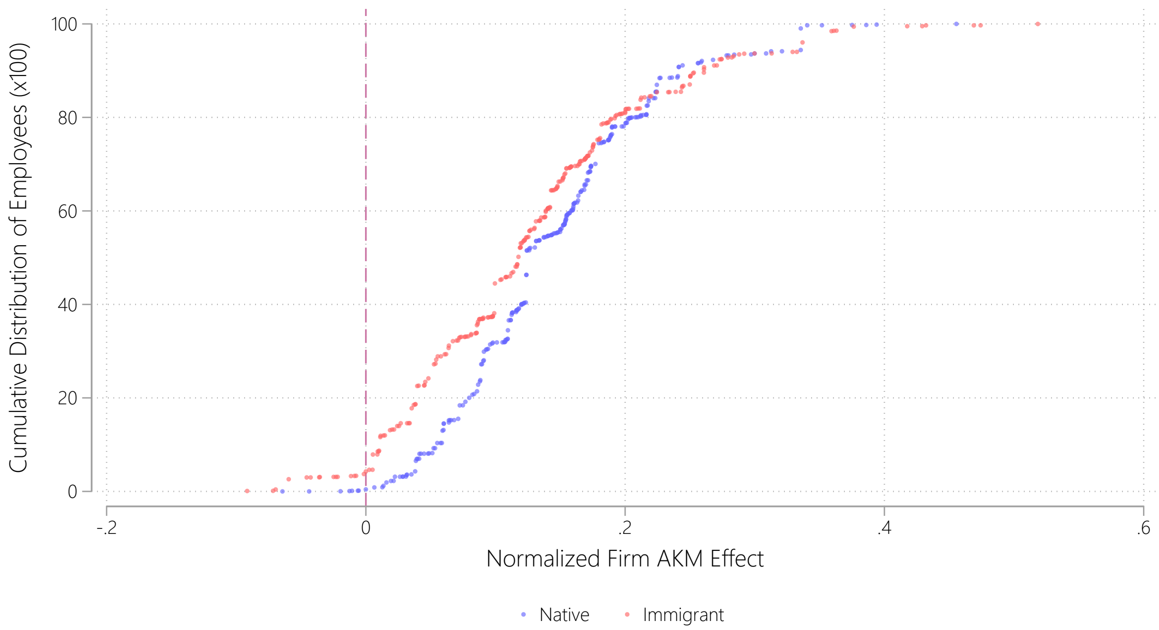
Education. Dostie et al. (2021) analyze the immigrant-native earnings gap by education but lack observations on natives' education. As can be seen in table 1 education can be observed for 75.5% (64.6%) of natives (immigrants) at least once in the dual-connected set. This allows an analysis of the immigrant-native earnings gap *within* the same education level.[32]
In the subgroup of tertiary-educated individuals the earnings gap turns negative, meaning immigrants' mean earnings are 4.0% higher than those of natives. Although this difference is mainly driven by 'transferable' skills, 17.5% of the gap can be attributed to firms. Even though the firm effect is relatively weak itself, it masks important dynamics. Well educated immigrants sort to high-premium firms better than natives, measured by a sorting effect of −3.2 ppts., this effect is almost completely offset by firm policies through the pay-setting effect of 2.5 ppts. To put this into context, the difference in transferable

[32]To my knowledge this is the first education-based decomposition of the immigrant-native gap. By the nature of the data, workers of big firms are more likely to have an observed education level. This type of selection would only pose a problem for the interpretability of the results if immigrants are *differently* selected than natives. The way the ESS and SE are conducted this seems unlikely.

worker characteristics accounts for 3.3 ppts. of the earnings gap. Therefore, for well educated workers, sorting and pay-setting effects individually are almost as important as worker productivity for the earnings gap.

The earnings gap for the subgroup of non-tertiary-educated individuals is 11.9%. The firm effect is especially high for this subgroup at 3.3 ppts. and accounts for 27.7% of the earnings gap, distributed relatively evenly over sorting and pay-setting effects, 15.1% and 12.6% respectively. Differences in time-invariant earnings capacity seem to be much more important for the earnings gap of non-tertiary-educated individuals than for tertiary-educated individuals, 95.0 vs. 22.5% contribution.

Figure 6
Cumulative Distribution of Normalized Firm AKM Effects by Group and 3-digit Industry



Notes: Cumulative distribution of normalized firm AKM effects by 3-digit industry. For each 3-digit industry the person-year weighted average normalized firm AKM effect (horizontal axis) and the number of person-year observations (vertical axis) is computed for both groups separately. Based on the dual-connected set.

The above observations seem to support the idea that the Swiss labor market is segmented. In the low-skill sector the firm effect seems to increase inequality, while this effect fades out the more skilled the sector is. This can be seen visually in figure 6. The distribution of firm AKM effects of natives and immigrants is clearly separated in low firm premium industries, whereas the distributions seem to converge, or even cross, with increasing firm premia.[33] The fade out is driven by the ability of immigrants to sort to high-premium firms better than natives, while a discrepancy between immigrants and

[33]The relative ranking of industries is not preserved across distributions of natives and immigrants.

natives wage regimes partly offsets this effect. Natives' steeper firm ladder can also be seen in the figure.

Firm Composition. CHK and Sorkin and Wallskog (2023) find that firm premium dispersion over time is largely driven by a changing composition of firms, i.e. newer cohorts of firms enter more dispersed and stay more dispersed throughout their lives.
Closely related to this is the question how much of the dynamics in sorting and pay-setting effects over time are driven by compositional changes in the pool of active firms. To analyze this aspect I employ the following decomposition of the change in sorting and pay-setting effects defined in equation 20 inspired by Haltiwanger (1997). Here, I only show the decomposition for the change in sorting effect, although the decomposition of the change in pay-setting effect is analogous and can be found together with the derivation of the decomposition in appendix C.2:

$$\begin{aligned}\Delta\bar{\psi}_t^{\text{sorting}} = \sum_j \psi_j^N(\Delta\pi_{Njt} - \Delta\pi_{Mjt}) = & \underbrace{\sum_{j\in\mathcal{J}_{t-1}\cap\mathcal{J}_t} \tilde{\psi}_{it}^N(\Delta\pi_{Njt} - \Delta\pi_{Mjt})}_{\Delta\text{ Re-sorting}} \\ & + \underbrace{\sum_{j\notin\mathcal{J}_{t-1}\wedge j\in\mathcal{J}_t} \tilde{\psi}_{jt}^N(\pi_{Njt} - \pi_{Mjt})}_{\Delta\text{ Firm Entry}} \\ & - \underbrace{\sum_{j\in\mathcal{J}_{t-1}\wedge j\notin\mathcal{J}_t} \tilde{\psi}_{jt}^N(\pi_{Njt-1} - \pi_{Mjt-1})}_{\Delta\text{ Firm Exit}},\end{aligned} \tag{28}$$

where $\mathcal{J}_t$ is the set of firms active in period t, $\tilde{\psi}_{jt}^g = \psi_j^g - \bar{\psi}_{t-1}^{\text{sorting}}$ and $\bar{\psi}_t^{\text{sorting}} = \sum_j \psi_j^N(\pi_{Njt} - \pi_{Mjt})$. The re-sorting effect gives the change in mean firm AKM effect of group g that is driven by reallocation of workers between firms staying active between $t-1$ and t, holding firm composition constant. The firm entry and exit components isolate the change originating from firms entering and exiting the labor market.
Figure A.10 shows the cumulative change in sorting and pay-setting effects decomposed using the above identity for all immigrants. Note that a *decrease* of the respective effect indicates a *more favorable* situation for immigrants. For both, sorting and pay-setting, changing firm composition leads to excess earnings growth of immigrants relative to natives. Entry of new firms accounts for a decrease of 0.4 ppts. of the sorting effect, while firm exit's effect on sorting is weak. New firms with a relatively high (low) firm premium seem to hire a higher (lower) share of immigrants than those firms already active or exiting the pool of firms.
Firm entry and exit is relatively more important for the pay-setting effect, where net

entry accounts for a decrease of 0.2 ppts. Re-sorting between active firms widens the immigrant-native earnings gap induced by unequal firm wage policies. Firms entering the pool of active firms have group-specific firm premia more in favor of immigrants than firms active on, or leaving the labor market.[34] Furthermore, net entry seems to persistently influence pay-setting over the whole sample period, while the effect stays relatively constant after the initial years in the sorting case.

Overall it can be concluded that firm wage policies are an important driver behind the immigrant-native earnings gap in Switzerland. The prominent role of pay-setting effects indicates that standard AKM models with common firm effects for immigrants and natives do not match the patterns found in the data. The decomposition results are similar the findings of Arellano-Bover and San (2024), who also estimate an important role of the pay-setting channel. Other literature analyzing group earnings gaps using AKM frameworks, namely Card, Cardoso, and Kline (2016), Gerard et al. (2021), Dostie et al. (2021), without exception measure a weak pay-setting effect and a strong sorting effect.

6.2 Earnings Assimilation over Time

Analysis Cohort. As shown by Borjas (1985), measuring earnings assimilation in the cross-section can lead to biased results due to compositional changes in the arriving cohorts. To measure assimilation in earnings and isolate this channel from the effect of skill differences between cohorts, the sample of immigrants is restricted to the cohort arriving in the five years around the sample's start, i.e. 2000–2004.[35] Column (7) of table 1 shows summary statistics for the cohort. Although there are some differences in terms of earnings, age, origin and education compared to all immigrants in the dual-connected set, these differences are not substantial. Further, immigrants are divided into two groups based on their country of origin. Immigrants from advantaged, economically well developed countries, namely Northern and Western Europe[36], the United States, Canada, Australia and New Zealand and immigrants from non-advantaged 'Other countries' including Southern Europe. Figure A.12 shows the average person AKM effects for natives and immigrants by origin × education subgroups and immigrants' distribution across these subgroups. Immigrants from economically well-developed countries are more often tertiary-educated, while immigrants from 'Other countries' are predominantly non-tertiary-educated. As expected, tertiary-educated immigrants generally have higher mean person AKM effects.

[34] Note that there is no mechanical reason for the entry effect to be greater in magnitude than the exit effect.

[35] The qualitative implications of the results are not sensible to minor shifts in the cohort time span.

[36] This includes Austria, Belgien, Denmark, Faroe Islands, Finland, France, Germany, Gibraltar, Iceland, Ireland, Liechtenstein, Luxembourg, Monaco, Netherlands, Norway, Sweden and the United Kingdom.

Furthermore, 'NW Eur./US/CA/AU/NZ' immigrants have higher mean person AKM effects than those from 'Other countries' or the respective native peer group.
Figure A.11 shows the number of immigrants and their composition in terms of origin by year of immigration for all immigrants in the dual-connected set.[37] As can be seen, most immigrants in the subset's time span emigrated from Western, Northern and Southern Europe. Immigrants from the Balkans form another big group. If there is earnings assimilation, it can be expected that mean earnings over the sample period are higher for the subsample than the sample of all immigrants in the dual-connected set, because cohorts immigrating later had less time to catch up. This could explain the higher earnings for the analysis cohort in table 1.
Figure A.11 also shows the mean person AKM effect by year of immigration. This measure shows the extent of compositional changes between cohorts over time. It can be clearly seen that unobserved worker 'skill' increases sharply from 1993 to 2000 after experiencing a dip in the late 1980s and early 1990s. The dip could be explained by the influx of relatively unskilled war refugees after the collapse of Yugoslavia. Afterwards, Switzerland experienced a shift towards highly educated immigrants (Dorn and Zweimüller, 2021), partially driven by skill-biased technological change increasing the relative demand for highly educated workers (Beerli, Indergand, and Kunz, 2023). Although there are some fluctuations after 2000, there is no trend evident, meaning that later cohorts are comparable in composition to the 2000–2004 analysis cohort.

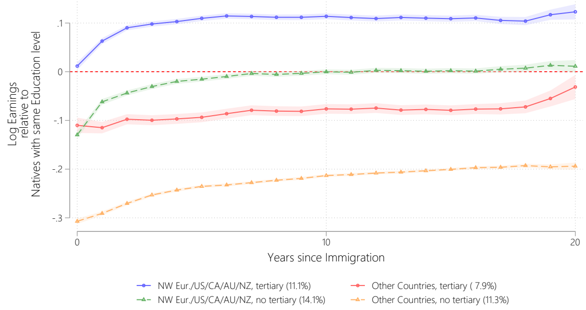
Assimilation. Figure 7(a) shows assimilation in log earnings relative to natives of the same education level by years since immigration for origin $\times$ education subgroups, i.e. estimates of $\{\beta_e\}_{e=0}^{20}$ from equation 24.[38] Standard errors are clustered at the person level. The cumulative change in log points is given in parentheses in the legend. Adjusted for compositional effects, well educated individuals experience an expected earnings growth of 7.9 ppts. if originating from 'Other countries', up to 11.1 ppts. for those from 'NW Eur./US/CA/AU/NZ'. Expected adjusted earnings of individuals lacking tertiary education increase by 11.3–14.1 ppts.
Most of earnings assimilation takes place within the first ten years after arrival. Well educated immigrants from 'Other countries' show a different pattern than the other subgroups. These individuals experience less excess growth in the first years, yet they show a higher rate of assimilation in the last three years.

[37] The high number of immigrants in 1981 might be related to the passing of the first Swiss asylum legislation in the same year. Also note that the figure only shows immigrants still working in Switzerland in the sample period.

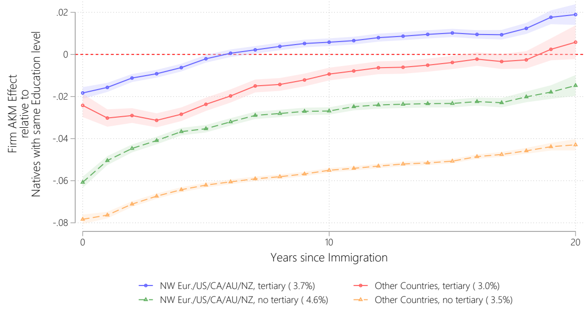
[38] As mentioned above, the equation is estimated by education subgroups, such that results give the earnings gap relative to natives of the *same* education level. Figure A.14 shows results comparing the subgroups to all natives irrespective of education. Qualitatively, results stay similar.

Figure 7
FIRM EFFECT ASSIMILATION

(a) LOG EARNINGS



(b) FIRM AKM EFFECT



Notes: Point estimates and 95% confidence intervals of $\{\beta_e\}_{e=0}^{20}$ from equations 21 and 24 for origin × education subgroups. Immigrants are compared to natives of the same education level. Standard errors are clustered at the individual level. Based on all natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort.

Figure 7(b) shows estimates of equation 21, i.e. immigrants' expected firm effect relative to natives of the same education level. In general, firm premia are higher for immigrants with tertiary education and those from economically well developed countries. Except for the tertiary-educated 'Other countries' subgroup, firm premia assimilation is steeper in the years right after immigration, then slows down and, for all subgroups, accelerates in the last years. Interestingly the firm effect contributes to earnings assimilation even after earnings assimilation has stopped for all subgroups, which means that it compensates other residual effects counteracting assimilation. The magnitude of the change in firm effect is stronger for less educated immigrants, 3.5–4.6 ppts., accounting for 31–32% of earnings catch up. Well educated immigrants show an positive change of 3.0–3.7 ppts, which explains 33–38% of assimilation in adjusted earnings. 'Other countries', non-tertiary-educated immigrants experience a dip in firm premia relative to natives, which explains the poor record in earnings assimilation for this subgroup. Tertiary-educated immigrants eventually catch up to natives' firm premia. Those originating from economically well developed countries do this a lot faster, 6 vs. 19 years, although starting from a similar level initially.

Figures 8(a)–8(d) decompose the firm effect into sorting and pay-setting contributions using equation 23. Rates of assimilation in the effects seem to be independent of origin and rather depend on education.[39]

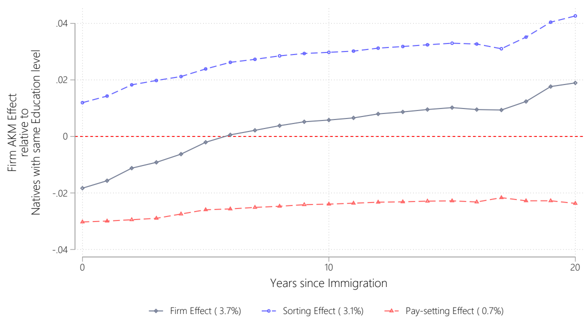
Parallel to the cross-sectional analysis in section 6.1, well-educated immigrants sort to high-premium firms 'better' than natives, even at arrival. This gives these immigrants

[39]Figure A.13 shows sorting and pay-setting effects across subgroups, making comparisons across subgroups easier.

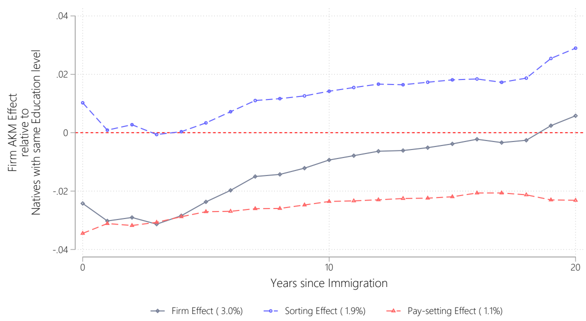
a lead of approximately 1 ppts. relative to natives right away. This effect is more than offset by within-firm wage disparities. The growth in firm premia is mainly driven by climbing the firm ladder defined by natives' premia. Changes in the pay-setting channel explain only about 19–37% of assimilation in firm premia, while sorting explains 63–81% for these subgroups.

Figure 8
Firm Effect Decomposition

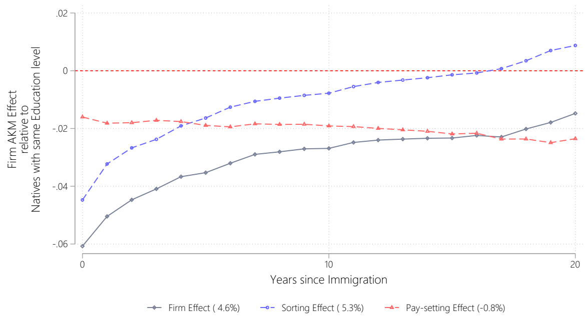
(a) NW Eur./US/CA/AU/NZ
tertiary



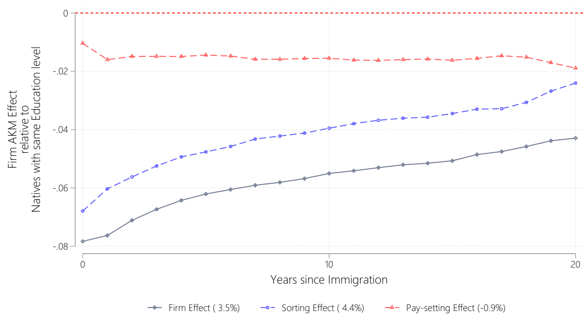
(b) Other Countries
tertiary



(c) NW Eur./US/CA/AU/NZ
no tertiary



(d) Other Countries
no tertiary



Notes: Point estimates of decomposition 23 for origin × education subgroups. Immigrants are compared to natives of the same education level. Based on all natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort.

Non-tertiary-educated immigrants initially allocate towards high-premium firms worse than comparable natives. On arrival, sorting explains approximately 70–90% of the firm AKM effect gap. Although these immigrants do climb the firm ladder to better-paying firms, a 4.4–5.3 ppts. increase, this is even faster than tertiary-educated immigrants, the destination firms have less favorable wage policies for immigrants relative to natives. Due to these less equitable firm policies the mean firm effect decreases by 0.8–0.9 ppts. relative to natives for these subgroups.

Overall, sorting to better paying firms, i.e. climbing the firm ladder, accounts for the majority of the change in firm effect for all subgroups, 63–126%, while gaps due to pay-setting are more persistent over time and show slow convergence or even a slight widening of the immigrant-native gap for non-tertiary-educated immigrants. As a result, all subgroups catch up in terms of firm AKM effects.

Figure B.3 shows estimates applied on the sample consisting of male workers only. Results are similar, with sorting explaining 51–92% of firm effect catch-up.

Selective Emigration. Selective emigration of individuals who earn less than ex ante expected naturally increases mean earnings of stayers and may give a false impression of assimilation (Borjas, 1985; Lubotsky, 2007).[40] To rule out this kind of 'survivorship bias' being a major driver behind the above assimilation results I follow a similar approach as Abramitzky, Boustan, and Eriksson (2014) and Dustmann and Görlach (2015) and restrict the cohort sample to individuals participating on the labor market until 2020. This is the case for 73.0% of the 2000–2004 cohort (see table 1).[41] Figure A.16 shows the results. Overall the results are very similar to the baseline results. This rules out systematic selective emigration as a driver of assimilation results.

Dostie et al. (2021) measure that mean person AKM effects of most immigrants *decrease* relative to Canadian natives. This fact is explained partly by selective emigration of the most successful earners mostly to the United States. In the Swiss context, there is no emigration destination with higher earnings potential in proximity, which could explain why this effect is weaker or absent in Switzerland.[42]

Age-at-Arrival. Immigrant age at immigration has been an important field of immigration research (see e.g. Friedberg (1992); Schaafsma and Sweetman (2001); Alexander and Ward (2018)). Accumulation of human capital is heavily dependent on the age of immigration. Schaafsma and Sweetman (2001) show that immigrants aged between 15 to 18 obtain the lowest years of education for Canadian data, while those immigrating with a younger age have comparable or even more years of schooling compared to natives. 1940 US census data show a similar pattern (Alexander and Ward, 2018). Figure 9 shows assimilation profiles by subgroups defined by age at arrival to Switzerland.

Interestingly, subgroups aged 10–29, i.e. teenagers and young adults at arrival, show the lowest levels of log earnings upon labor market entrance combined with slower con-

[40] For models of selective return migration see Yezer and Thurston (1976) and Dustmann and Görlach (2015).

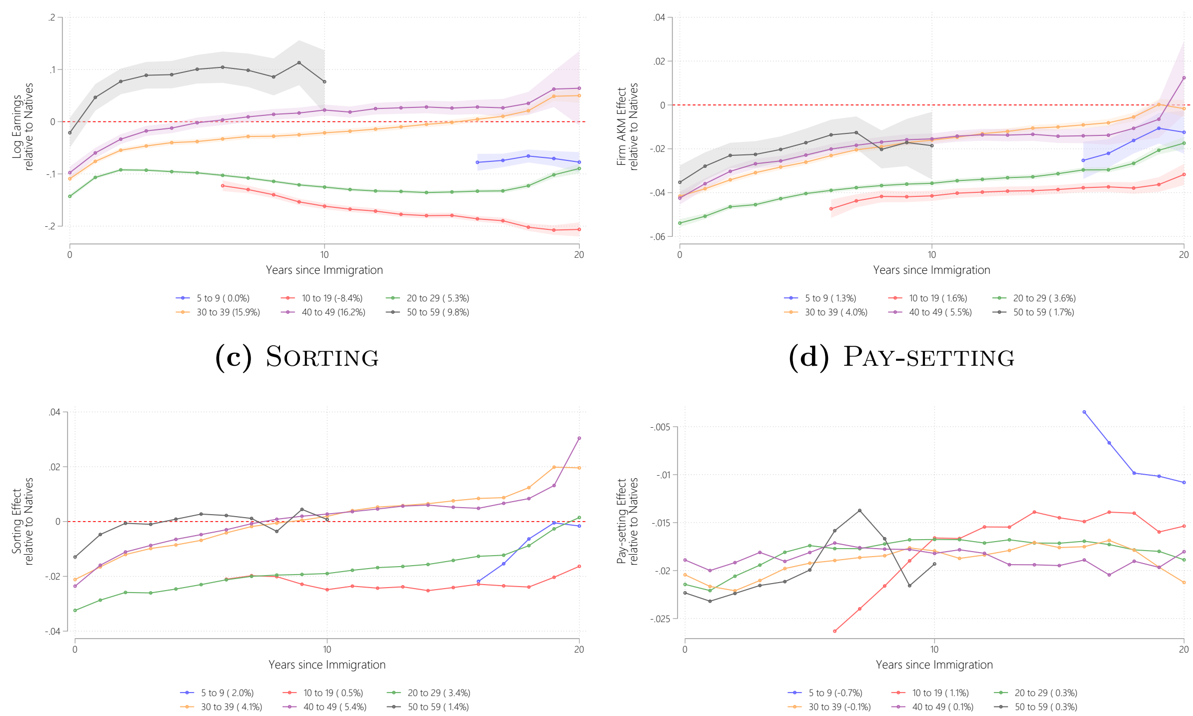
[41] This share is partially driven down by individuals immigrating with high age who naturally do not participate on the labor market for long due to old age.

[42] Figure A.9 shows mean person AKM effects for natives and immigrants by years for *all* migrants similar to Dostie et al. (2021). Although all subgroups' mean person AKM effect decreases, natives experience the sharpest decline. The decline in expected time-invariant earnings capacity can be attributed to many factors. These include the rise in female participation (see Federal Statistical Office (2023b)) and part-time work (see Federal Statistical Office (2023c)) which could be relatively more pronounced for natives.

vergence towards natives or even a widening gap. Looking at firm effect estimates, the gap in log earnings can be, at least partially, explained by lower firm premia. In turn, virtually all of the firm premia gap is can be attributed to these immigrants sorting to well-paying firms worse than their younger and older peers. Pay-setting discrimination is rather homogeneous across subgroups. This is consistent with the findings of Schaafsma and Sweetman (2001), who find that earnings are lowest for immigrants arriving in their late teens. This can be explained by this age being a crucial point in life, entering a new society and culture leads to these immigrants acquiring less schooling. Another effect that could explain this: potential immigrants still attending college are more likely to wait until they finish their degree before migrating. In most countries, students are between 18 and 25 years old. This could lead to an adverse selection of immigrants of this age at arrival.

Figure 9

Assimilation - by Age at Arrival



(a) Log Earnings

(b) Firm AKM Effect

(c) Sorting

(d) Pay-setting

Notes: Point estimates and 95% confidence intervals of $\{\beta_e\}_{e=0}^{20}$ from equations 21 and 24 and point estimates of decomposition 23 for age-at-arrival subgroups. Immigrants are compared to all natives. Standard errors are clustered at the individual level. Based on all natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort.

Figure A.18(a) shows mean person AKM effects by age at arrival for the analysis cohort. Time-invariant earnings capacity visibly decreases between ages 11–23, supporting both of the above hypotheses. Furthermore, when restricting the sample to immigrants lacking tertiary education, see figure A.18(b), immigrants aged 10–29 at arrival show less adverse firm premium outcomes. Consistent with the idea of adverse selection, it is important to note that age at arrival is correlated with unobservable worker characteristics influencing earnings. For example, immigrants coming with the age of 50 are much more likely to be in a managerial position than those coming with the age of 30. This explains the high log earnings and firm premia observations for the older immigrants. Again, when looking at estimates for low-educated immigrants, these vanish. Immigrants arriving between ages 5–9 seem to form an exception, these show only a small gap in firm effects, which quickly decreases, this is likely due to more and earlier exposure to local culture.

Overall, age at immigration does not seem to directly influence firm effect assimilation beyond individuals arriving as young children or the indirect effects induced by less schooling when arriving as teenagers or young adults.

Arellano-Bover and San (2024) do not find differences in firm premium assimilation across age-at-arrival nor an adverse outcome for individuals immigrating as teenagers or young adults.

Cohorts. As mentioned above, the analysis was conducted using the cohort of immigrants arriving between 2000 and 2004. If either the capability of the Swiss labor market or the composition of immigrants has changed towards a state where assimilation over time takes place faster or is less necessary due to better performance on arrival this should be evident in the data. Instead of following the approach by Abramitzky, Boustan, and Eriksson (2014) and including a cohort fixed effect, allowing for cohort-specific intercepts in the outcome variables, I additionally allow for heterogeneous slopes like Arellano-Bover and San (2024). This enables the model to estimate different rates of convergence for each cohort.

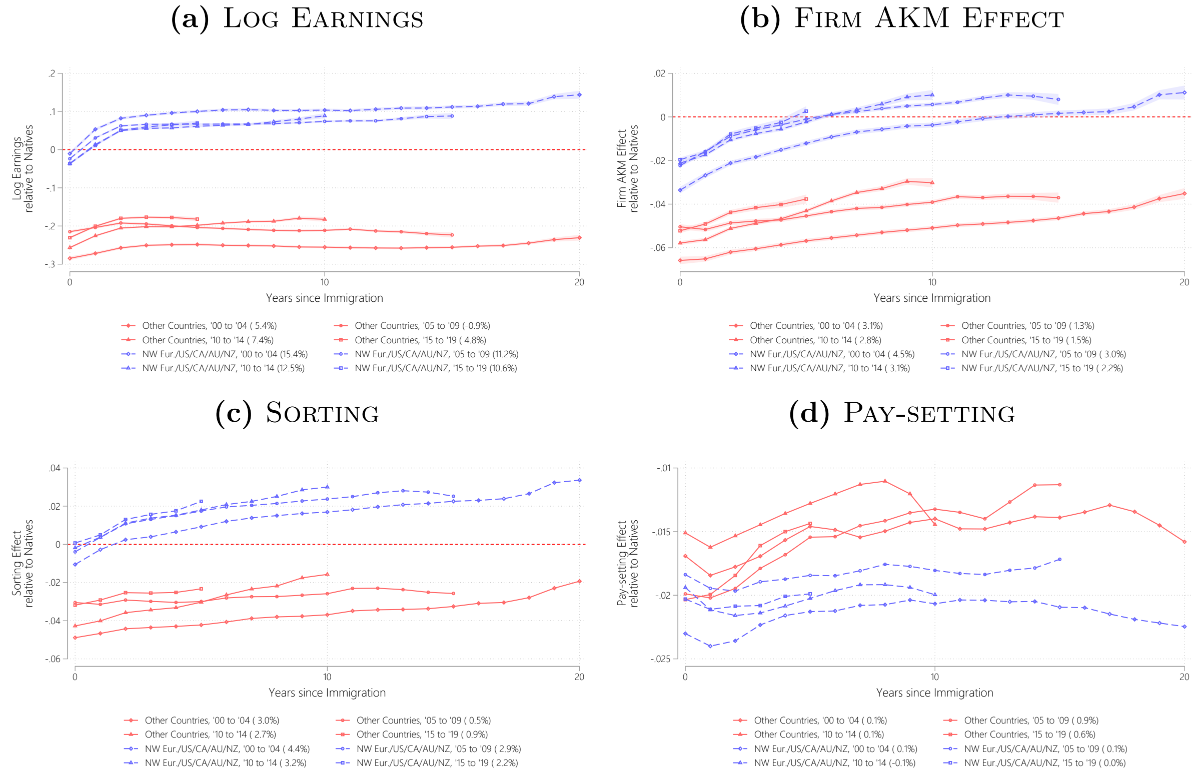
Figure 10 shows estimates of equations 24 and 23 by origin × cohort subgroups, where immigrants are compared to all natives. Cohort estimates may convolute cohort effects with compositional changes in skill-levels between cohorts. To mitigate this concern I estimate cohort effects by a more detailed analysis utilizing origin × cohort × education subgroups, these are presented in figure A.17. Reassuringly, the patterns are similar to the origin × cohort analysis, ruling out compositional effects as a driver of cohort effects. For reasons of visual clarity the figure is not presented in the main text.

Later cohorts earn, on average, higher average firm premia at arrival, although the inter-cohort differences can get really small as for 'NW Eur./US/CA/AU/NZ'-immigrants. This is almost entirely driven by better sorting, while allocation to firms with more equitable pay regimes shows no quantitatively big differences between cohorts. The effect is more pronounced for immigrants from economically less developed countries where there is a difference of approximately 2 ppts. between the ′00–′04 and ′15–′19 cohorts. Apart from this, assimilation paths are roughly parallel. The Swiss labor market has improved its ability to allocate new immigrants to high premium firms.

The 'Other countries' cohort arriving between 2005–2009 does not quite follow the pattern described above, as it has higher initial earnings and firm premia than its predecessors and successors. This is mostly driven by well-educated 'Other countries'-immigrants showing intertwined assimilation paths between cohorts as can be seen in figure A.17. Arellano-Bover and San (2024) find similar cohort patterns.

Figure 10

ASSIMILATION - BY COHORT × ORIGIN



Notes: Point estimates and 95% confidence intervals of $\{\beta_e\}_{e=0}^{20}$ from equations 21 and 24 and point estimates of decomposition 23 for cohort × origin subgroups. Immigrants are compared to all natives. Standard errors are clustered at the individual level. Based on all natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort.

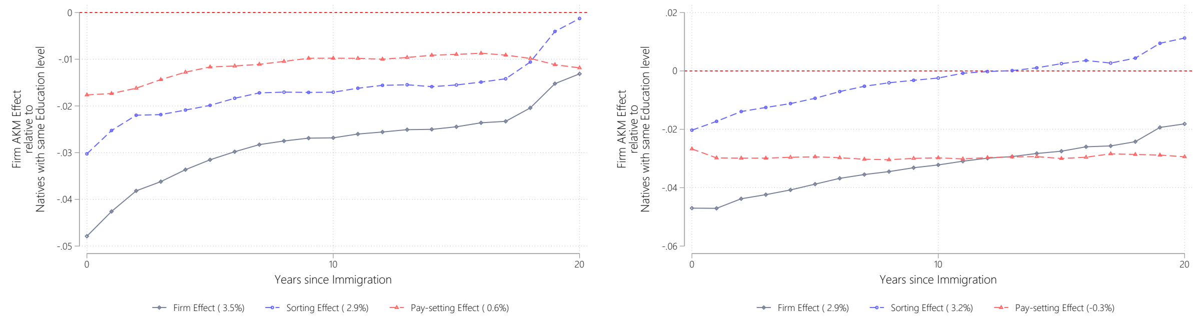
Gender. Figure 11 shows the decomposition of firm effect gaps into sorting and pay-setting by gender. Each gender is compared to natives of the same gender. Men and women face comparable firm effect gaps, albeit the drivers behind these differ. While the men's gap is mainly driven by sorting which explains approximately 63% of the initial gap, the women's gap is primarily driven by pay-setting which explains approximately 54% of the gap on arrival.

For both, most of firm effect assimilation is driven by mobility towards firms with generally higher wages (82% for men, 110% for women), while the pay-setting effect is relatively persistent over time. This catch-up takes place predominantly in the first years and again after 18 years for males. Females show a much more linear assimilation process.

Figure 11

Assimilation - by Gender

(a) Male **(b)** Female



Notes: Point estimates of decomposition 23 for gender subgroups. Immigrants are compared to natives of the same gender. Based on all natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort.

Culture. Switzerland is unique in that it has four different official languages, alongside the languages also cultural attitudes change. National referendums repeatedly revealed strong differences in political attitudes and preferences between the 'Latin' and 'German' parts. For example 'Latin'-speaking parts of Switzerland have been historically more supportive for referendums demanding less weekly working hours or longer vacations (Eugster et al., 2017). As shown by Eugster et al. (2017), these cultural differences lead to higher unemployment durations for 'Latin' than German-speaking individuals, even within the same jurisdictions and similar labor markets.

This might imply that cultural differences or similarities might hinder respectively ease assimilation in firm premia.

Figure 12 shows estimates of firm AKM effect gaps relative to natives in the respective regions and the decomposition into sorting and pay-setting by years since arrival for tertiary-educated immigrants from Germany, Austria and France to German- and French-speaking regions in Switzerland.[43]

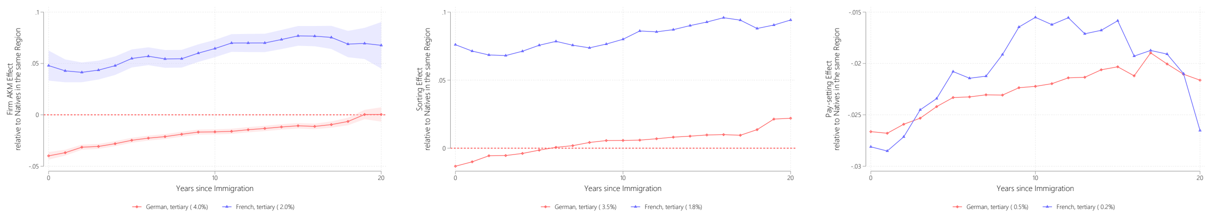
German-speaking immigrants in French-speaking regions demand a much higher mean firm AKM effect than those working in the German-speaking part of Switzerland and vice-versa. This self-selection could be driven by adjustment and opportunity costs the immigrant has to bear in order to adapt to a different culture and language. For the German-speaking regions, the higher firm premia for French Immigrants are predominantly driven by better sorting to high-paying firms. There is no evidence for systematic differences in pay-setting, i.e. French immigrants are not discriminated more or less rela-

[43]Due to the low number of observations I restrain from showing estimates for Italian- or Rhaeto-Romance-speaking regions.

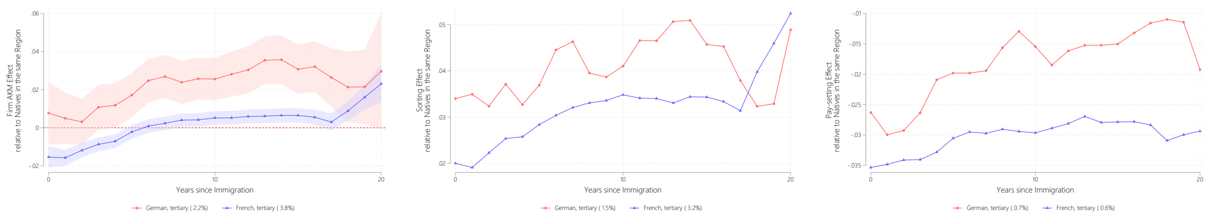
tive to German and Austrian Immigrants.
On the other hand, German and Austrian immigrants to French-speaking regions are less impacted by unequal firm wage policies.
Immigrants where the origin culture corresponds to the destination culture experience higher firm premia growth than 'cross-immigrants', 4.0 and 3.9% vs. 2.0 and 2.2%. Though this could be explained due to the already higher initial earnings level of 'cross-immigrants', which leaves less potential for earnings growth.
Figure A.19 shows estimates for non-tertiary-educated immigrants. Here the adjustment cost effect vanishes for French immigrants to German-speaking regions. Immigrants show a similar assimilation path irrespective of country of origin. Restricting the sample to immigrants still present in 2020 does not change the results.
Overall, there is no conclusive evidence for heterogeneous assimilation of immigrants of similar or different cultures, this leaves room for further research into this topic.

Figure 12
Assimilation - by Cultural Regions

(a) German-speaking



(b) French-speaking



Notes: Point estimates and 95% confidence intervals of $\{\beta_e\}_{e=0}^{20}$ from equation 21 and point estimates of decomposition 23 for tertiary-educated German, Austrian and French immigrants in German- and French-speaking labor market regions. Immigrants are compared to tertiary-educated natives residing in the defined regions. Standard errors are clustered at the individual level. Based on natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort who reside in the respective German- or French-speaking regions. German-speaking regions are: Bern, Basel, Aarau-Olten, Zurich, Winterthur-Schaffhausen, St. Gallen, Chur, Lucerne. French-speaking regions are: Geneva, Lausanne, Sion, Fribourg, Neuchâtel.

7 Convergence Mechanisms

In this section I follow Arellano-Bover and San (2024) and estimate equation 21 using different outcomes in order to further analyze immigrants' behavior and differences between subgroups after arrival to Switzerland.

7.1 Econometric Framework

In a model of time-invariant firm premia, climbing the firm ladder requires a change of employer and the subsequent employer to offer a higher firm premium (Arellano-Bover and San, 2024). To measure job mobility of immigrants relative to natives I estimate the following regression:

$$\mathbb{1}\{j(igt) \neq j(igt-1)\} = \mathbb{1}\{g(i) = M\} \times \left[\sum_{e=0}^{20} \beta_e \mathbb{1}\{E_{it} = e\} \right] + X_{it}\theta + \varepsilon_{it}. \tag{29}$$

To measure the magnitude of the steps on the firm ladder, I condition the sample to job movers, i.e. $\mathbb{1}\{j(igt) \neq j(igt-1)\} = 1$, and estimate the following regression:

$$\hat{\psi}^g_{j(igt)} - \hat{\psi}^g_{j(igt-1)} = \mathbb{1}\{g(i) = M\} \times \left[\sum_{e=0}^{20} \beta_e \mathbb{1}\{E_{it} = e\} \right] + X_{it}\theta + \varepsilon_{it}. \tag{30}$$

To measure how immigrants assimilate towards firms which are more likely to have a native workforce, I estimate the following equation:

$$\rho_{ij(igt)t} = \mathbb{1}\{g(i) = M\} \times \left[\sum_{e=0}^{20} \beta_e \mathbb{1}\{E_{it} = e\} \right] + X_{it}\theta + \varepsilon_{it}, \tag{31}$$

where $\rho_{ij(igt)t} \equiv \frac{\sum_k \mathbb{1}\{j(kgt)=j(igt) \ \wedge \ g(k)=M \ \wedge \ k \neq i\}}{\sum_k \mathbb{1}\{j(kgt)=j(igt) \ \wedge \ k \neq i\}}$ is the share of immigrant coworkers at firm j in year t.

Last, I estimate the following equation to measure assimilation in terms of firm size:

$$\ln FirmSize_{ij(igt)t} = \mathbb{1}\{g(i) = M\} \times \left[\sum_{e=0}^{20} \beta_e \mathbb{1}\{E_{it} = e\} \right] + X_{it}\theta + \varepsilon_{it}, \tag{32}$$

where $FirmSize_{ij(igt)t} \equiv \sum_k \mathbb{1}\{j(kgt) = j(igt)\}$ is the natural logarithm of the number of employees of individual i's firm $j(igt) = j$. The above equations can be estimated for different subgroups of natives and immigrants, such as origin × education.

7.2 Results.

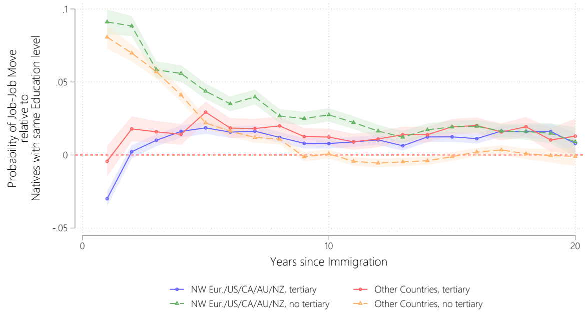
Figure 13 shows estimates of equations 29, 30, 31 and 32 by origin × education subgroups. Estimates compare immigrants with natives of the *same* education level.[44] Standard er-

[44] Figure A.14 shows results comparing the subgroups to all natives irrespective of education. Qualitatively results stay similar.

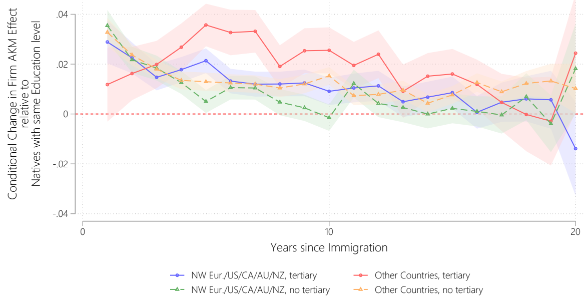
rors are clustered at the person level. The relevant natives' baseline is given in the notes below the respective figure.

Figure 13
ASSIMILATION - ADDITIONAL OUTCOMES

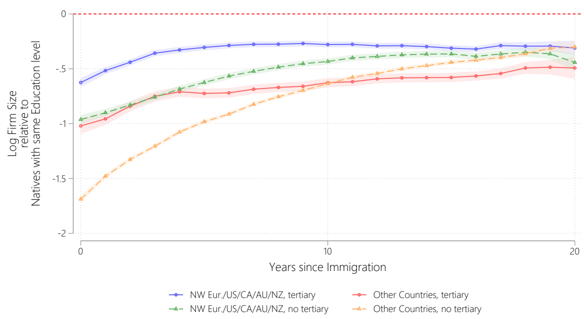
(a) JOB-JOB CHANGE PROBABILITY



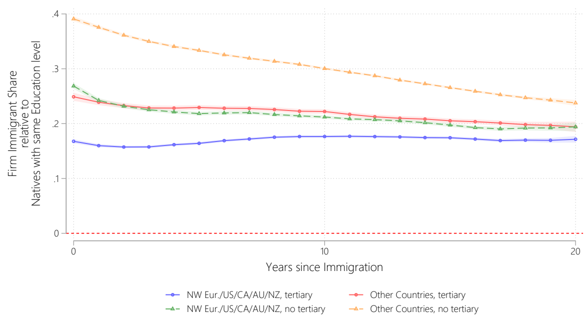
(b) GAIN IN FIRM PREMIUM CONDITIONAL ON JOB CHANGE



(c) FIRM SIZE



(d) COWORKER IMMIGRANT SHARE



Notes: Point estimates and 95% confidence intervals of $\{\beta_e\}_{e=0}^{20}$ from equations 29, 30, 31 and 32 by origin × education subgroups. Immigrants are compared to natives of the same education level. Standard errors are clustered at the individual level. Based on all natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort. Natives' baseline moments:
$\mathbb{E}[\mathbb{1}\{j(igt) \neq j(igt-1)\} \mid N, tertiary] = 0.11$
$\mathbb{E}[\mathbb{1}\{j(igt) \neq j(igt-1)\} \mid N, non-tertiary] = 0.10$
$\mathbb{E}[\hat{\psi}^g_{j(igt)} - \hat{\psi}^g_{j(igt-1)} \mid N, tertiary, switcher] = 0.012$
$\mathbb{E}[\hat{\psi}^g_{j(igt)} - \hat{\psi}^g_{j(igt-1)} \mid N, non-tertiary, switcher] = 0.0079$
$\mathbb{E}[\rho_{ij(igt)t} \mid N, tertiary] = 0.26$
$\mathbb{E}[\rho_{ij(igt)t} \mid N, non-tertiary] = 0.28$

Job Search. The trajectories of employer change probabilities in figure 13(a) depend heavily on the individual's education level, while showing only small differences by origin. Immigrants lacking tertiary education are approximately 7–9 ppts., i.e. 70–90%, more likely than their native peers to change their employer in the first year after immigration. This excess probability almost completely vanishes within 10 years after immigration and stabilizes between 0–2 ppts. Contrary to this, well educated immigrants are initially less likely to change employer than natives. The excess probability increases within the first

5 years to 2–3 ppts. or 18–27% above natives, whereafter it stabilizes at 1–2 ppts.

Step Size. Figure 13(b) shows that, conditional on an employer change, immigrants' firm premia increase by 3–4 ppts. more than those of natives in the first year after arrival. Relative to the native mean conditional step of 0.0079 and 0.012 for non-tertiary and tertiary natives respectively, these immigrants take 2.5–4.5 times larger steps on the firm ladder. Tertiary-educated 'Other countries' immigrants show a lower point estimate initially, explaining the dip seen in figure 7(b), but their trajectory stays higher for a longer time. With time, excess firm premia jump height fades out. In general, 'Other countries' immigrants show higher excess firm premium jumps.
Figure A.15 shows gains in firm premia not conditioning on employer changes, i.e. the interaction of the probability to take a step and the expected size of the step on the firm ladder. The higher gains of tertiary-educated immigrants from well developed countries is partly offset due to their lower probability of employer changes. Therefore the overall effect is similar for both origin groups for the first 4 years after arrival. Arellano-Bover and San (2024) estimate quantitatively similar profiles of employer change probabilities and gains in firm premia.

Firm Size and Immigrant Coworkers. It has been long established in the literature that there is a positive relationship between firm size and wage levels which cannot be easily explained by observable firm or worker characteristics, see e.g. Brown and Medoff (1989). As shown by Winter-Ebmer and Zweimüller (1999) this relationship holds in Switzerland and the differential cannot be fully explained by employee heterogeneity. Therefore firm size is of interest as an outcome. Immigrants' origin is an important factor in assimilation regarding firm size as can be seen in figure 13(c). Well educated immigrants and immigrants from well developed countries in general start with a higher firm size than other immigrants. Nevertheless, their employer is still smaller than the average native's firm. The differences are sizeable, and range from approximately 82% (1.7 log points) to 45% (0.6 log points) smaller firms than comparable natives. With the years, all immigrants increase their expected firm size, those starting with a lower level show a faster rate of growth. As a result all subgroups show an expected firm size approximately 3–5% smaller than their native peers after 20 years.
The share of immigrant coworkers behaves similarly, depicted in figure 13(d). Although there is virtually no assimilation for tertiary-educated immigrants from economically developed countries measured by this outcome. The immigrant share settles at approximately 18–24 ppts. above that of the respective native comparison group. Natives' baseline coworker immigrant share lies between 26–28%.

Appendix D presents additional results for the trajectory of managerial responsibilities and occupation rank by years since immigration.

7.3 Additional Analysis

Age-at-Arrival Figure A.20 shows estimates by age-at-arrival subgroups. Parallel to the lower realized firm premia documented in section 6.2, immigrants arriving as teenagers or young adults show lower job mobility than other immigrants. Arriving in these crucial years for human capital accumulation seems to make this subgroup face problems in employer search and change.
Immigrants arriving as children show outcomes very similar or indistinguishable to natives. The Swiss institutions are successful at integrating immigrants arriving in years of early development.

Cohorts. Figure A.21 shows estimates of equations 29, 30, 31 and 32 by origin × cohort subgroups.[45] Differences between cohorts in employer change probabilities and firm ladder steps conditional on employer changes are especially evident for immigrants from 'Other countries'. Later cohorts of this origin generally show higher employer change probabilities and larger firm ladder steps than their predecessors. Later cohorts of immigrants from economically well developed countries are less mobile initially, but their mobility seems to be elevated more consistently, manifested by a smaller slope.
In terms of firm size, later cohorts work at bigger firms after arrival while simultaneously showing lower rates of assimilation. As a result all cohorts seem to converge towards a similar level of firm size relative to natives. There are only weak cohort effects concerning the coworker immigrant share outcome.

Gender. Figure A.22 shows estimates comparing male and female immigrants relative to natives of the same gender. Although native men and women have the same employer change probabilities, immigrant women show less job mobility than immigrant men. Men are persistently 1–2 ppts. more likely to transition from job to job. This corresponds to approximately 10–20% of the native baseline. Furthermore, men stay more mobile even after 20 years, while female immigrants show virtually the same transition probability as natives after 11 years.
Conditioning on a job change, women make roughly 4 and men roughly 2.8 times larger steps on the firm ladder in the first year compared to natives. The effect quickly decreases

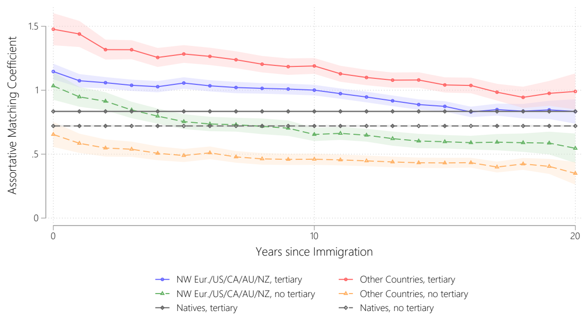
[45] Figure A.17 shows estimates by origin × cohort × education to rule out compositional differences between cohorts as a driver of heterogeneous trajectories.

for men, while women show a more persistent excess gain over time.
The interaction of both, mobility and size of steps, can be seen in figure A.22(c). Arellano-Bover and San (2024) find a similar behavior for male and female immigrants.
Figure A.22(d) shows that women work in firms with a higher share of natives on arrival and throughout the trajectory.

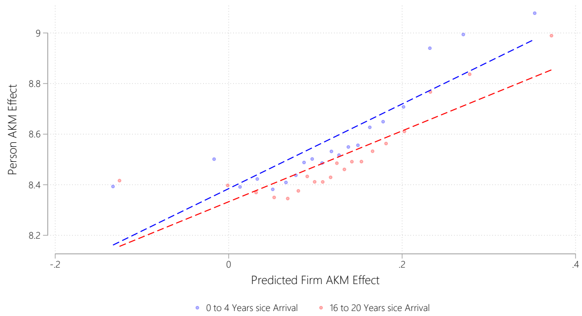
Assortative Matching. In order to understand how assortative matching changes over time, I estimate equation 27 separately by years since migration and origin × education subgroups using 2SLS. Figure 14(a) shows the estimates. Additionally, natives' cross-sectional estimates are displayed as horizontal lines.
Except for one subgroup, immigrants show higher assortative matching than their native peers on arrival. Tertiary-educated immigrants are in general more sorted. For all subgroups assortative matching decreases with time since immigration.
Arellano-Bover and San (2024) find initial sorting to be lower than for natives and a sharp increase in assortative matching in the first years after arrival for former USSR immigrants to Israel, consistent 'with low search capital and the hasty acceptance of whatever job was available on arrival'.

Figure 14
Assimilation - Assortative Matching

(a) Years since Migration



(b) Cross-section



Notes: Panel (a) shows point estimates and 95% confidence intervals of $\hat{\gamma}_1$ from equation 27 by years since arrival by origin × education subgroups. Cross-sectional estimates for natives by education are indicated by horizontal lines. Panel (b) shows a binned scatter plot and an OLS-fitted line of person AKM effects vs. predicted firm AKM effects separately for 0–4 and 16–20 years since arrival. Predicted firm AKM effects are obtained using natives' firm AKM effects as predictors. This is done to reduce attenuation bias due to estimation errors. Heteroskedasticity-robust standard errors are used. Based on all natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort.

To understand the driver behind the declining coefficients in the Swiss data, figure 14(b) shows a binned scatter plot for all immigrants in the first and last five years since arrival in the analysis cohort. The prediction of firm AKM effects based on native firm AKM effects is on the x-axis, while the y-axis shows person AKM effects. Visibly, the slope in the last five years is smaller, parallel to figure 14(a). The fitted line shifts to the right, meaning that, on average, all immigrants increased their firm premia. High person AKM effect-immigrants realize a larger relative increase than those on the lower tail of the distribution, this effect decreases the assortative matching coefficient.[46]

8 Conclusion

This paper used uniquely rich administrative data to analyze the influence of firm-specific wage policies on the immigrant-native earnings gap and earnings assimilation of immigrants in Switzerland. Using an AKM-style model I estimate that heterogeneous firm premia explain 22.1% of the earnings gap in the cross-section. 16.8% of this can be attributed to lower within-firm pay regimes for immigrants than for natives. 5.3% can be attributed to natives being allocated 'better' to high-premium firms. Estimating these by education levels reveals nuanced effects.

The inability of firms to perfectly discriminate between natives and immigrants can have many effects, for example it enables monopsony power over immigrants to 'spill over' to the wages of natives as in Amior and Manning (2023) or immigration to encourage firms to drop to and decrease wages at the bottom of the offer distribution as in Amior and Stuhler (2022). The strong role of differential pay-setting, controlling for education, shows that firms have the ability to discriminate between workers to a large extent, even *within* the same skill level. This is an important insight for the above mentioned models of monopsonistic wage setting. The adverse effects of immigration onto natives' wages could therefore be alleviated in the Swiss context.

Channels of immigrant convergence towards natives have been a much researched topic for a long time (see e.g. Chiswick (1978)). Recent literature has identified the 'firm-ladder'-channel as an important driver, it explains 31–38% of total growth in earnings within 20 years since arrival. While in the cross-section the pay-setting effect is more prominent, virtually all assimilation in firm premia over time is driven by immigrants sorting to, in general, higher paying firm as opposed to switching to firms with a lower wage penalty for immigrants. The source of within-firm wage differences and their persistence provides a promising area for future research.

In light of recent political debates in developed economies, the economic success of immigrants is not only of intrinsic interest for immigrants themselves, but simultaneously reduces the burden on social welfare systems, potentially increasing the acceptance of immigration as a solution to, among others, demographic problems.

[46]This effect is also evident when plotting figure 14(b) by origin $\times$ education subgroups.

References

Abowd, John M, Robert H Creecy, Francis Kramarz et al. 2002. "Computing person and firm effects using linked longitudinal employer-employee data." Tech. rep., Center for Economic Studies, US Census Bureau.

Abowd, John M, Francis Kramarz, Paul Lengermann, and Sébastien Pérez-Duarte. 2004. "Are good workers employed by good firms? A test of a simple assortative matching model for France and the United States." *Unpublished manuscript* 5.

Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica* 67 (2):251–334.

Abowd, John M, Kevin L McKinney, and Ian M Schmutte. 2019. "Modeling endogenous mobility in earnings determination." *Journal of Business & Economic Statistics* 37 (3):405–418.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2014. "A nation of immigrants: Assimilation and economic outcomes in the age of mass migration." *Journal of Political Economy* 122 (3):467–506.

Alexander, Rohan and Zachary Ward. 2018. "Age at arrival and assimilation during the age of mass migration." *The Journal of Economic History* 78 (3):904–937.

Alvarez, Jorge, Felipe Benguria, Niklas Engbom, and Christian Moser. 2018. "Firms and the decline in earnings inequality in Brazil." *American Economic Journal: Macroeconomics* 10 (1):149–189.

Amior, Michael and Alan Manning. 2023. "Monopsony and the wage effects of migration." Tech. rep.

Amior, Michael and Jan Stuhler. 2022. "Immigration and monopsony: Evidence across the distribution of firms." *mimeo* .

Andrews, Martyn J, Len Gill, Thorsten Schank, and Richard Upward. 2008. "High wage workers and low wage firms: negative assortative matching or limited mobility bias?" *Journal of the Royal Statistical Society Series A: Statistics in Society* 171 (3):673–697.

Andrews, Martyn J, Leonard Gill, Thorsten Schank, and Richard Upward. 2012. "High wage workers match with high wage firms: Clear evidence of the effects of limited mobility bias." *Economics Letters* 117 (3):824–827.

Arellano-Bover, Jaime and Shmuel San. 2024. "The Role of Firms and Job Mobility in the Assimilation of Immigrants: Former Soviet Union Jews in Israel 1990-2019." .

Bagger, Jesper and Rasmus Lentz. 2019. “An empirical model of wage dispersion with sorting.” *The Review of Economic Studies* 86 (1):153–190.

Bassier, Ihsaan, Arindrajit Dube, and Suresh Naidu. 2022. “Monopsony in movers: The elasticity of labor supply to firm wage policies.” *Journal of Human Resources* 57 (S):S50–s86.

Beerli, Andreas, Ronald Indergand, and Johannes S Kunz. 2023. “The supply of foreign talent: how skill-biased technology drives the location choice and skills of new immigrants.” *Journal of Population Economics* 36 (2):681–718.

Beerli, Andreas, Jan Ruffner, Michael Siegenthaler, and Giovanni Peri. 2021. “The abolition of immigration restrictions and the performance of firms and workers: evidence from Switzerland.” *American Economic Review* 111 (3):976–1012.

Berger, Marius and Bruno Lanz. 2020. “Minimum wage regulation in Switzerland: survey evidence for restaurants in the canton of Neuchâtel.” *Swiss Journal of Economics and Statistics* 156:1–23.

Bonhomme, Stéphane, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad, and Bradley Setzler. 2023. “How much should we trust estimates of firm effects and worker sorting?” *Journal of Labor Economics* 41 (2):291–322.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa. 2019. “A distributional framework for matched employer employee data.” *Econometrica* 87 (3):699–739.

Borjas, George J. 1985. “Assimilation, changes in cohort quality, and the earnings of immigrants.” *Journal of labor Economics* 3 (4):463–489.

Brown, Charles and James Medoff. 1989. “The employer size-wage effect.” *Journal of Political Economy* 97 (5):1027–1059.

Bruns, Benjamin. 2019. “Changes in workplace heterogeneity and how they widen the gender wage gap.” *American Economic Journal: Applied Economics* 11 (2):74–113.

Card, David. 2022. “Who set your wage?” *American Economic Review* 112 (4):1075–1090.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. 2018. “Firms and labor market inequality: Evidence and some theory.” *Journal of Labor Economics* 36 (S1):S13–S70.

Card, David, Ana Rute Cardoso, and Patrick Kline. 2016. “Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women.” *The Quarterly journal of economics* 131 (2):633–686.

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

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

Coudin, Elise, Sophie Maillard, and Maxime Tô. 2018. "Family, firms and the gender wage gap in France." Tech. rep., IFS Working Papers.

Damas de Matos, Ana. 2017. "Firm heterogeneity and immigrant wage assimilation." *Applied Economics Letters* 24 (9):653–657.

Dorn, David and Josef Zweimüller. 2021. "Migration and labor market integration in Europe." *Journal of Economic Perspectives* 35 (2):49–76.

Dostie, Benoit, Jiang Li, David Card, and Daniel Parent. 2021. "Employer policies and the immigrant–native earnings gap." *Journal of econometrics* .

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*, vol. 1. Elsevier, 489–533.

Dustmann, Christian, Hyejin Ku, and Tetyana Surovtseva. 2024. "Real exchange rates and the earnings of immigrants." *The Economic Journal* 134 (657):271–294.

Eeckhout, Jan and Philipp Kircher. 2011. "Identifying sorting—in theory." *The Review of Economic Studies* 78 (3):872–906.

Engbom, Niklas and Christian Moser. 2022. "Earnings inequality and the minimum wage: Evidence from Brazil." *American Economic Review* 112 (12):3803–3847.

Eugster, Beatrix, Rafael Lalive, Andreas Steinhauer, and Josef Zweimüller. 2017. "Culture, work attitudes, and job search: Evidence from the Swiss language border." *Journal of the European Economic Association* 15 (5):1056–1100.

Federal Statistical Office, FSO. 2000. "Die 16 Arbeitsmarktregionen der Schweiz am 5.12.2000 (Volkszählung)." URL https://www.bfs.admin.ch/asset/de/465644.

———. 2022a. "Arbeitsmarktregionen und Arbeitsmarktgrossregionen." URL https://www.bfs.admin.ch/bfs/en/home/statistics/territory-environment/nomenclatures/lma.assetdetail.21024457.html.

———. 2022b. “Bevölkerung nach Migrationsstatus.” URL https://www.bfs.admin.ch/bfs/de/home/statistiken/bevoelkerung/migration-integration/nach-migrationsstatuts.html#:~:text=Im%20Jahr%202022%20hatten%202,Personen)%20besass%20die%20Schweizer%20Staatsangeh%C3%B6rigkeit.

———. 2023a. “Arbeitsproduktivität nach Branchen zu laufenden Preisen.” URL https://www.experimental.bfs.admin.ch/expstat/de/home/projekte/arbeitsproduktivitaet-branchen.assetdetail.26985266.html.

———. 2023b. “Erwerbsbeteiligung.” URL https://www.bfs.admin.ch/bfs/de/home/statistiken/wirtschaftliche-soziale-situation-bevoelkerung/gleichstellung-frau-mann/erwerbstaetigkeit/erwerbsbeteiligung.html.

———. 2023c. “Teilzeitarbeit.” URL https://www.bfs.admin.ch/bfs/de/home/statistiken/wirtschaftliche-soziale-situation-bevoelkerung/gleichstellung-frau-mann/erwerbstaetigkeit/teilzeitarbeit.html.

Fox, Jeremy T. 2009. “Firm-size wage gaps, job responsibility, and hierarchical matching.” *Journal of Labor Economics* 27 (1):83–126.

Friedberg, Rachel M. 1992. “The labor market assimilation of immigrants in the United States: the role of age at arrival.” *Brown University* .

Gallen, Yana, Rune V Lesner, and Rune Vejlin. 2019. “The labor market gender gap in Denmark: Sorting out the past 30 years.” *Labour Economics* 56:58–67.

Gerard, François, Lorenzo Lagos, Edson Severnini, and David Card. 2021. “Assortative matching or exclusionary hiring? the impact of employment and pay policies on racial wage differences in brazil.” *American Economic Review* 111 (10):3418–3457.

Gibbs, Michael, Kathryn Ierulli, and Eva Meyersson Milgrom. 2003. “Careers in firm and occupational labor markets.” Tech. rep.

Goldschmidt, Deborah and Johannes F Schmieder. 2017. “The rise of domestic outsourcing and the evolution of the German wage structure.” *The Quarterly Journal of Economics* 132 (3):1165–1217.

Gruetter, Max and Rafael Lalive. 2009. “The importance of firms in wage determination.” *Labour Economics* 16 (2):149–160.

Haltiwanger, John. 1997. “Measuring and analyzing aggregate fluctuations: the importance of building from microeconomic evidence.” *Federal Reserve Bank of St. Louis Review* (May):55–78.

Iranzo, Susana, Fabiano Schivardi, and Elisa Tosetti. 2008. "Skill dispersion and firm productivity: An analysis with employer-employee matched data." *Journal of Labor Economics* 26 (2):247–285.

Jochmans, Koen and Martin Weidner. 2019. "Fixed-effect regressions on network data." *Econometrica* 87 (5):1543–1560.

Kleven, Henrik, Camille Landais, and Gabriel Leite-Mariante. 2023. "The child penalty atlas." Tech. rep., National Bureau of Economic Research.

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten. 2020. "Leave-out estimation of variance components." *Econometrica* 88 (5):1859–1898.

Lachowska, Marta, Alexandre Mas, Raffaele Saggio, and Stephen A Woodbury. 2023. "Do firm effects drift? Evidence from Washington administrative data." *Journal of Econometrics* 233 (2):375–395.

Lalive, Rafael and Tobias Lehmann. 2020. "The labor market in Switzerland, 2000-2018." *IZA World of Labor* .

Lopes de Melo, Rafael. 2018. "Firm wage differentials and labor market sorting: Reconciling theory and evidence." *Journal of Political Economy* 126 (1):313–346.

Lubotsky, Darren. 2007. "Chutes or ladders? A longitudinal analysis of immigrant earnings." *Journal of Political Economy* 115 (5):820–867.

Macis, Mario and Fabiano Schivardi. 2016. "Exports and wages: Rent sharing, workforce composition, or returns to skills?" *Journal of Labor Economics* 34 (4):945–978.

Malchow-Møller, Nikolaj, Jakob R Munch, and Jan Rose Skaksen. 2012. "Do immigrants affect firm-specific wages?" *The Scandinavian Journal of Economics* 114 (4):1267–1295.

McFadden, Daniel et al. 1973. "Conditional logit analysis of qualitative choice behavior." .

Naidu, Suresh, Yaw Nyarko, and Shing-Yi Wang. 2016. "Monopsony power in migrant labor markets: evidence from the United Arab Emirates." *Journal of Political Economy* 124 (6):1735–1792.

Oaxaca, Ronald. 1973. "Male-female wage differentials in urban labor markets." *International economic review* :693–709.

Rho, Deborah and Seth Sanders. 2021. "Immigrant earnings assimilation in the United States: a panel analysis." *Journal of Labor Economics* 39 (1):37–78.

Schaafsma, Joseph and Arthur Sweetman. 2001. "Immigrant earnings: age at immigration matters." *Canadian journal of Economics* :1066–1099.

Schmidheiny, Kurt and Michaela Slotwinski. 2018. "Tax-induced mobility: Evidence from a foreigners' tax scheme in Switzerland." *Journal of Public Economics* 167:293–324.

Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter. 2019. "Firming up inequality." *The Quarterly journal of economics* 134 (1):1–50.

Sorkin, Isaac. 2017. "The role of firms in gender earnings inequality: Evidence from the United States." *American Economic Review* 107 (5):384–387.

———. 2018. "Ranking firms using revealed preference." *The quarterly journal of economics* 133 (3):1331–1393.

Sorkin, Isaac and Melanie Wallskog. 2023. "The Slow Diffusion of Earnings Inequality." Working Paper 30977, National Bureau of Economic Research.

White, Halbert. 1980. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* 48 (4):817–838.

Winter-Ebmer, Rudolf and Josef Zweimüller. 1999. "Firm-size wage differentials in Switzerland: Evidence from job-changers." *American Economic Review* 89 (2):89–93.

Yezer, Anthony MJ and Lawrence Thurston. 1976. "Migration patterns and income change: implications for the human capital approach to migration." *Southern Economic Journal* :693–702.

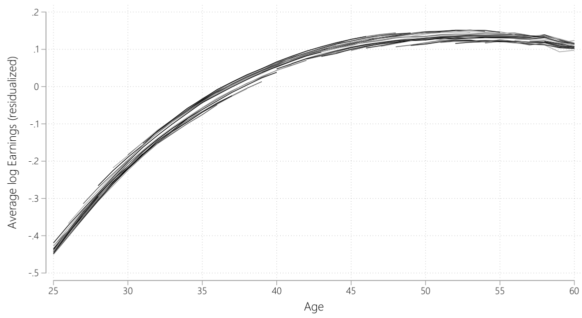
Appendices

A Additional Figures and Tables

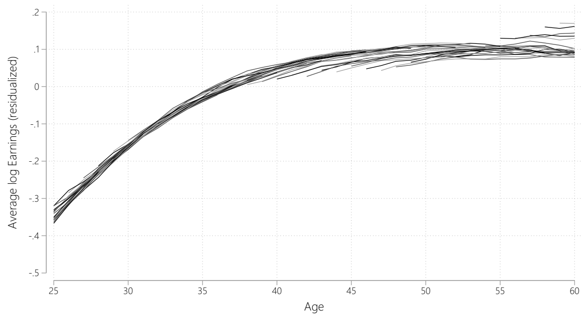
A.1 Figures

Figure A.1
Earnings-Age Profile

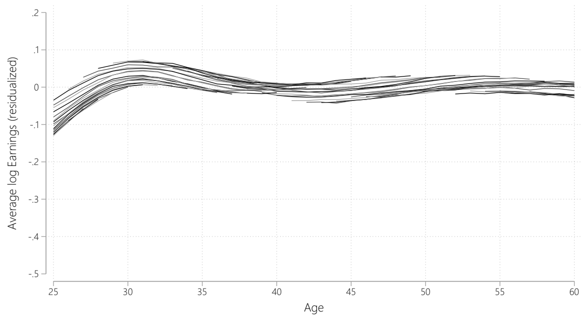
(a) Swiss Male



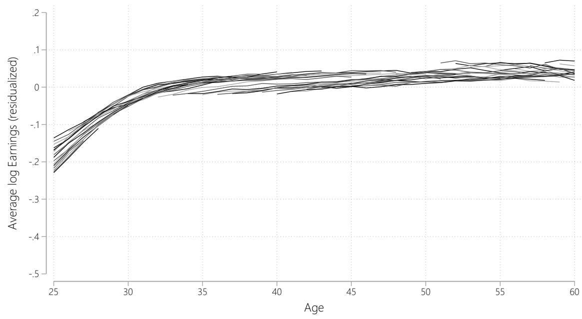
(b) Immigrant Male



(c) Swiss Female

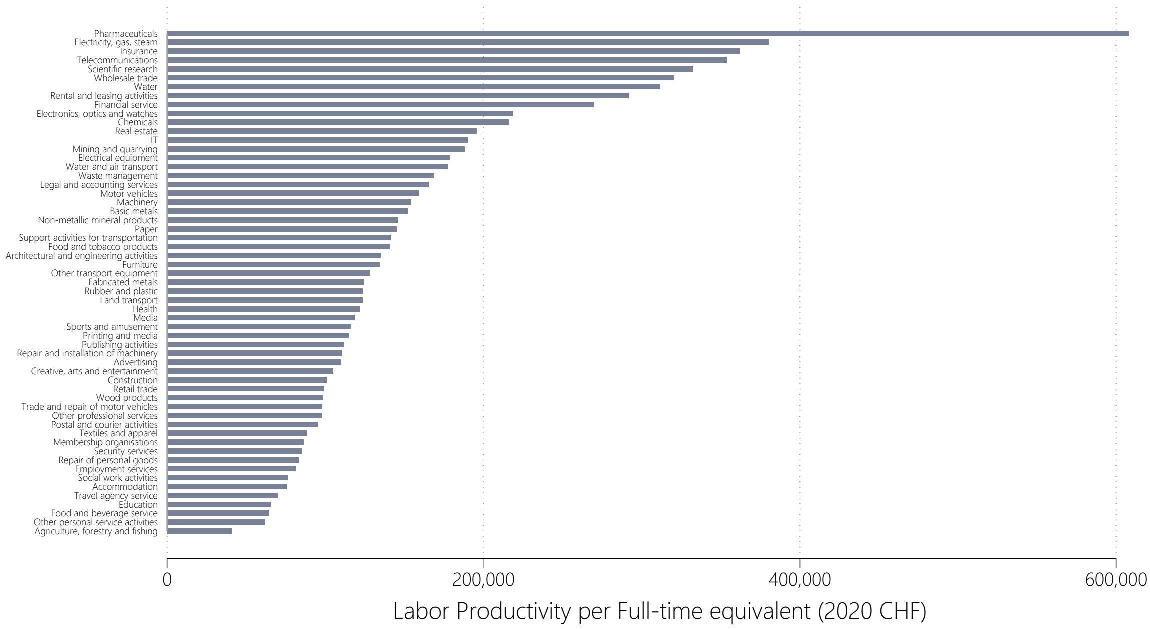


(d) Immigrant Female



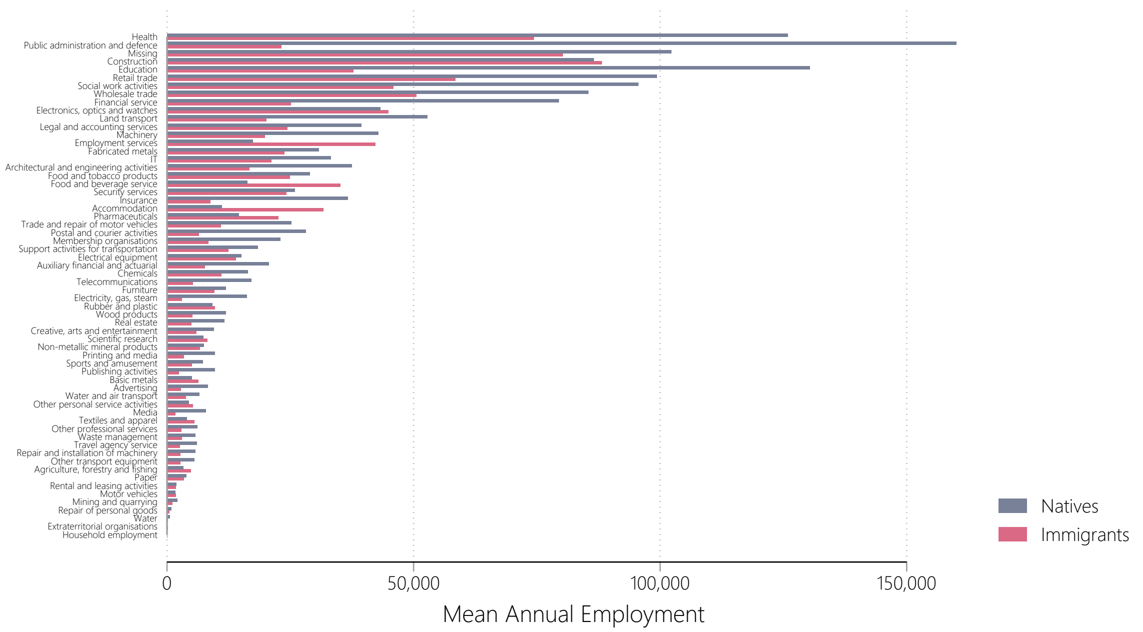
Notes: Each line shows the age-earnings profile of one specific cohort of the respective origin-gender group. The profiles are obtained by residualizing mean log wages of each cohort and age using year fixed effects. Based on the largest connected set for each origin group. Figure inspired by Gerard et al. (2021).

Figure A.2
LABOR PRODUCTIVITY BY 2-DIGIT INDUSTRY



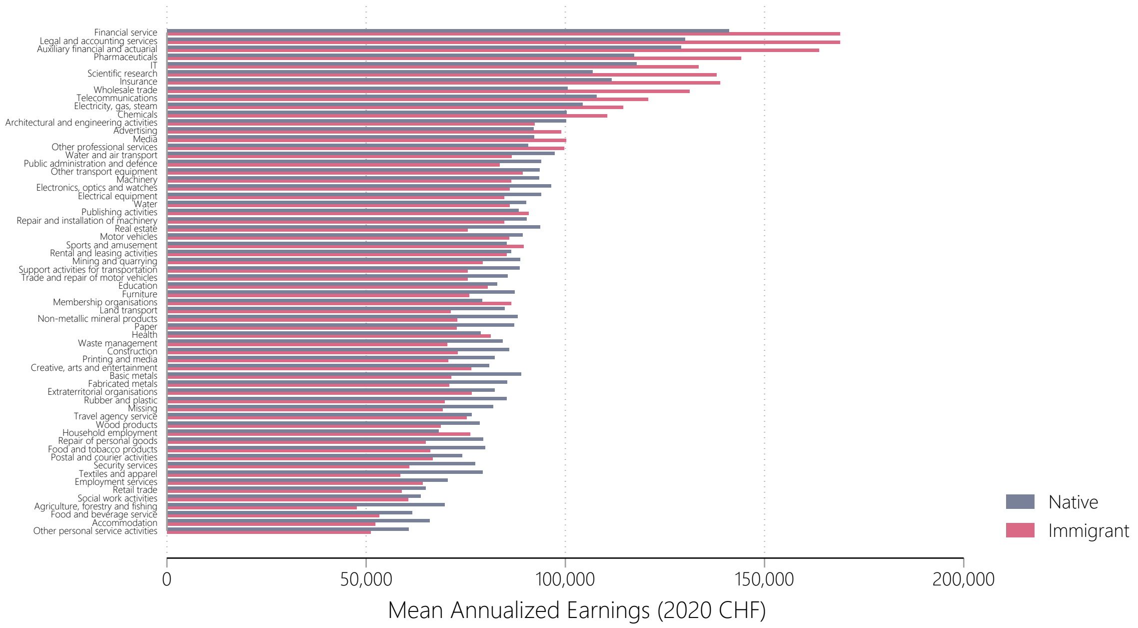
Notes: The figure shows the labor productivity per full-time equivalent by 2-digit NOGA industries. The data are based on the Federal Statistical Office (2023a), adjusted to 2020 CHF and averaged over the years 2002–2020.

Figure A.3
EMPLOYMENT BY 2-DIGIT INDUSTRY



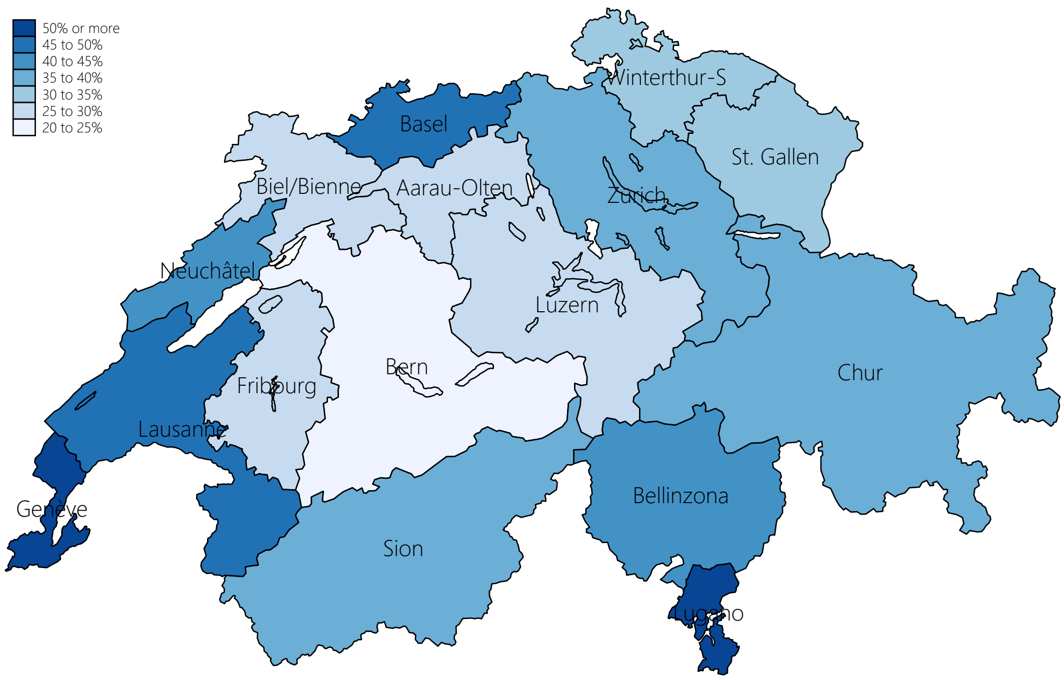
Notes: Mean annual employment by 2-digit industry and immigrant status. Based on the dual-connected set.

Figure A.4
Earnings by 2-digit Industry



Notes: Mean annualized earnings by 2-digit industry and immigrant status. Based on the dual-connected set.

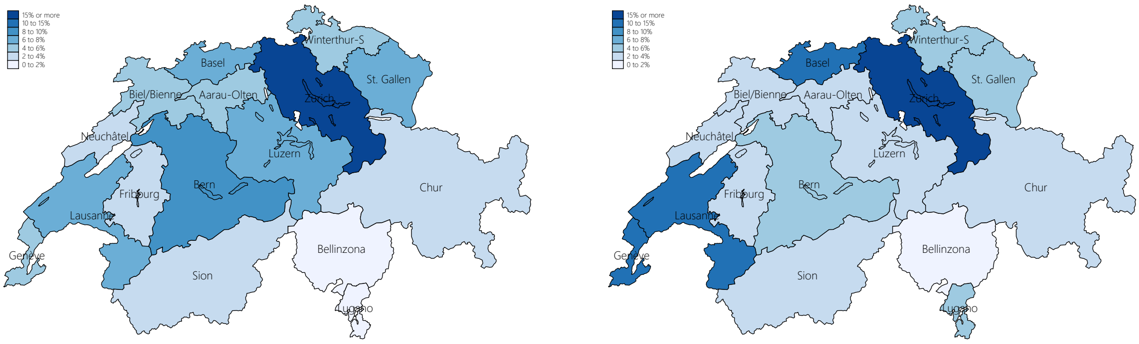
Figure A.5
Immigrant Share by Labor Market Regions



Notes: The figure shows the share of immigrant person-year observations by labor market regions. Based on the dual-connected set.

Figure A.6
SPATIAL DISTRIBUTION BY LABOR MARKET REGIONS

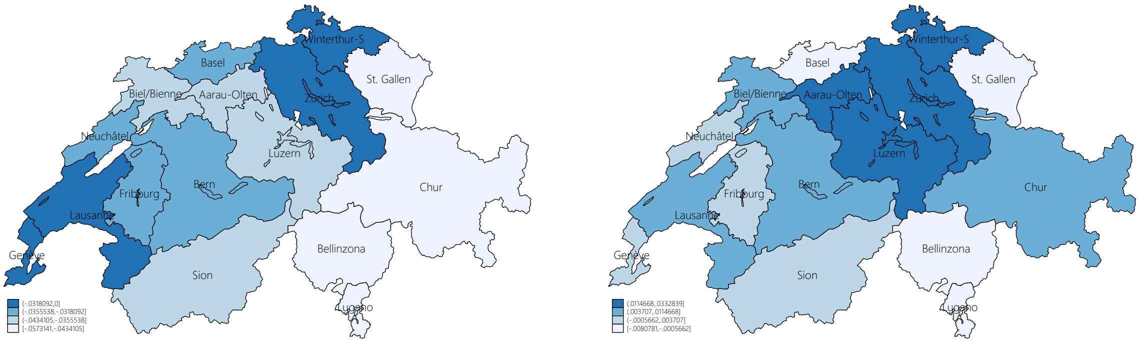
(a) NATIVES **(b)** IMMIGRANTS



Notes: Panel (a) shows the distribution of native person-year observations over labor market regions. Panel (b) shows the same distribution for immigrants. Based on the dual-connected set.

Figure A.7
AKM REGION FIXED EFFECT ESTIMATES

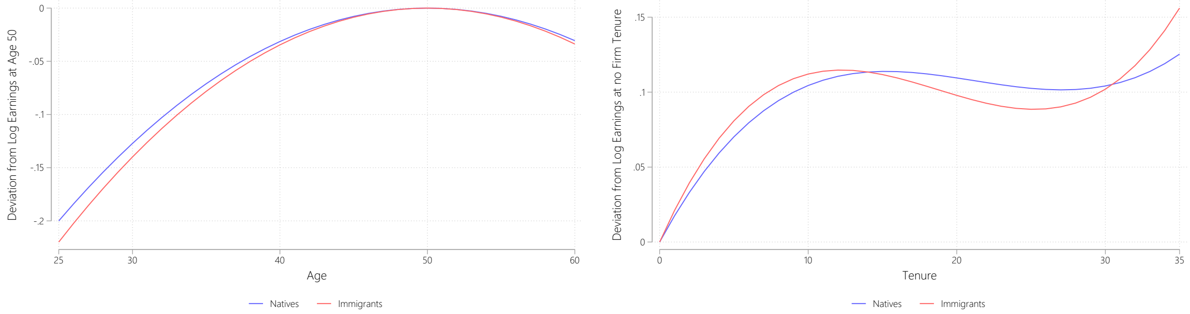
(a) NATIVES **(b)** IMMIGRANTS



Notes: The figure shows estimates of the labor market region fixed effects for natives and immigrants. Based on the largest connected set for each origin group.

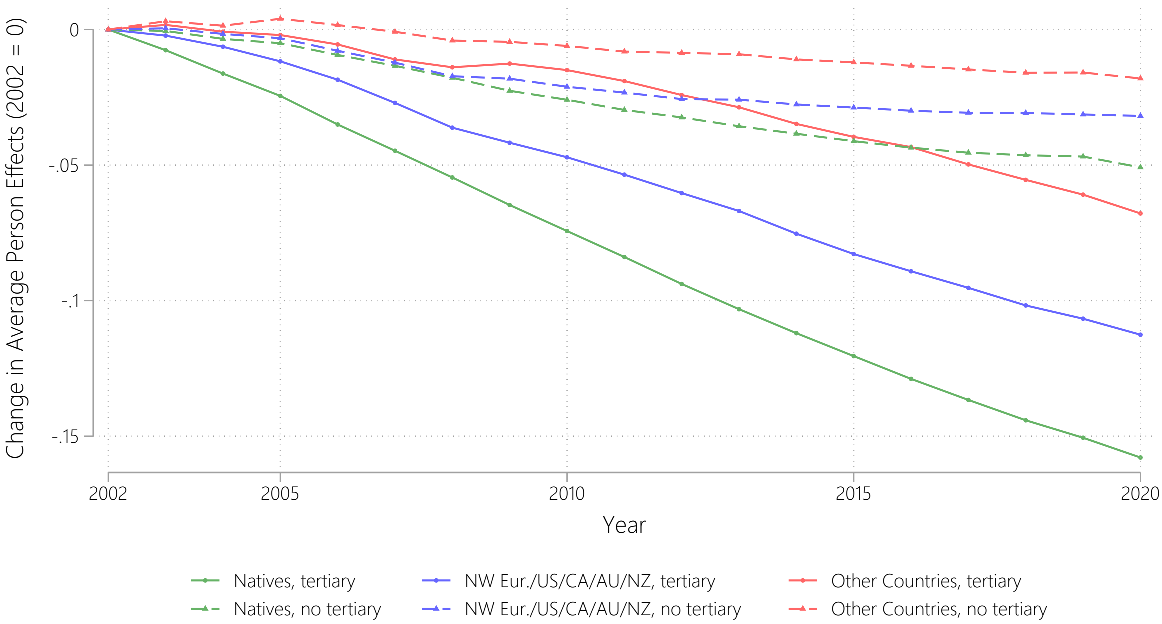
Figure A.8
ESTIMATED AGE AND TENURE PROFILES

(a) AGE **(b)** TENURE



Notes: Estimates of age and tenure profiles from the AKM model. Based on the largest connected set for each origin group.

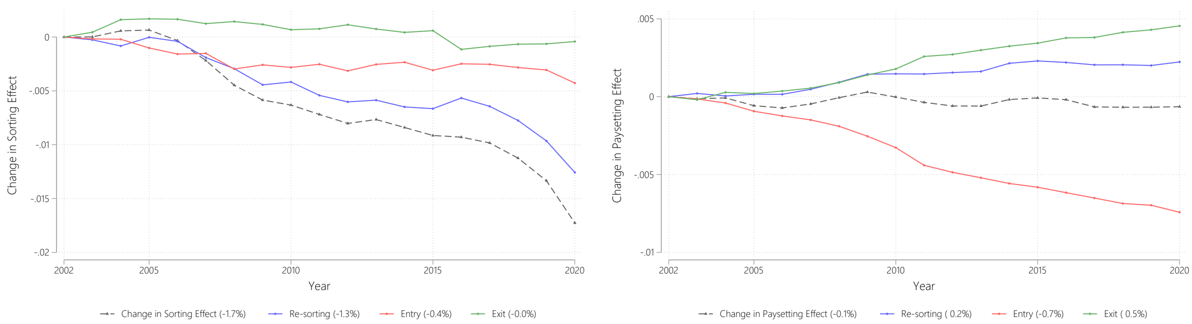
Figure A.9
PERSON AKM EFFECTS OVER TIME



Notes: The figure shows the change of mean person effects relative to 2002 for origin × education subgroups. Based on the dual-connected set.

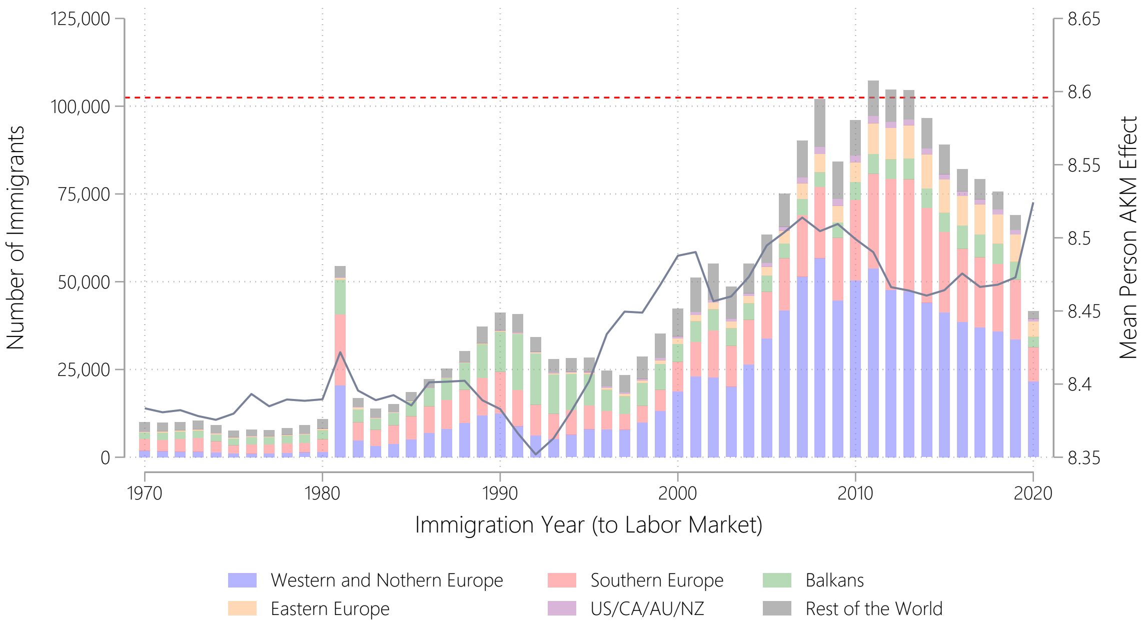
Figure A.10
FIRM COMPOSITION CONTRIBUTION TO FIRM EFFECT

(a) SORTING **(b)** PAY-SETTING



Notes: Estimates of decomposition 28. For visual reasons, the components are cumulated over time. All immigrants are compared to all natives. Based on the dual-connected set.

Figure A.11
Immigrant Composition by Year of Arrival

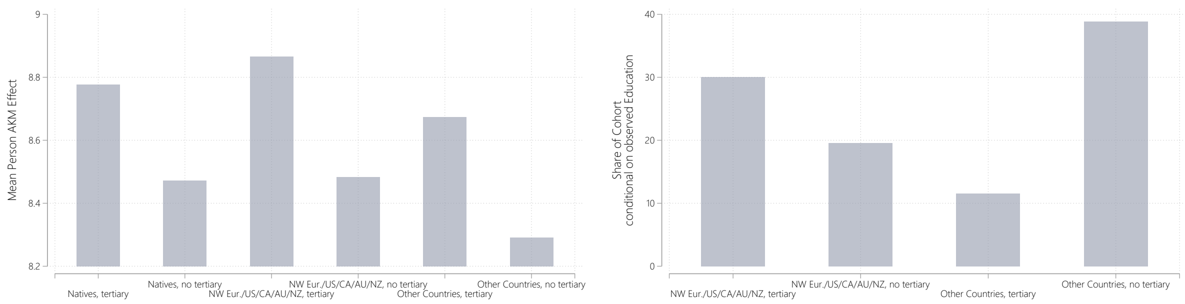


Notes: The figure shows the total number, origin and mean person AKM effect of immigrants by year of arrival. Only immigrants with observed immigration year are taken into account. The immigration year of CBW's is set to the first year of appearance in the overall sample. Numbers for immigrants before 1970 are not displayed. The dashed red horizontal line indicates the mean natives person AKM effect. Based on the dual-connected set.

Figure A.12
2000–2004 Cohort Composition

(a) Person AKM Effect

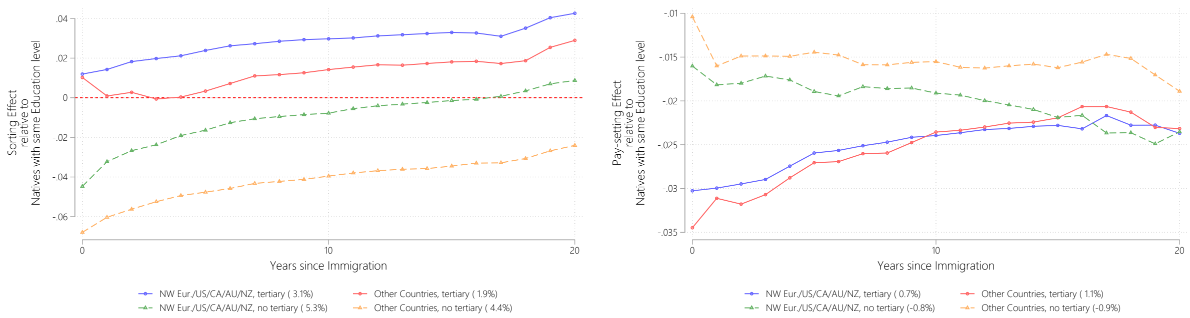
(b) Shares



Notes: Panel (a) shows mean person AKM effects for natives and immigrants by origin × education subgroups. Panel (b) shows the distribution of immigrants to origin × education subgroups conditioning on observed education. Based on all natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort.

Figure A.13
SORTING AND PAY-SETTING ACROSS SUBGROUPS

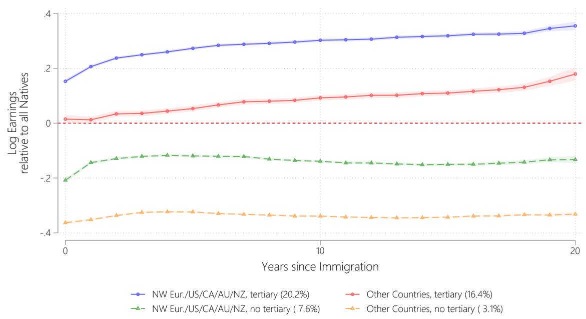
(a) SORTING **(b)** PAY-SETTING



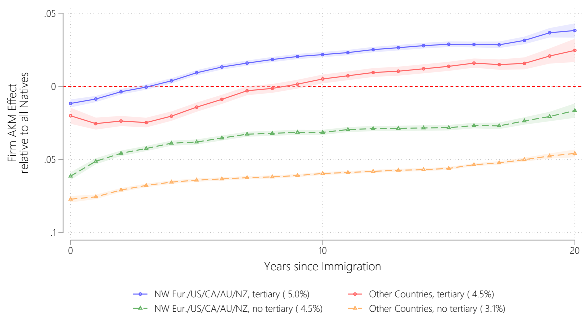
Notes: Point estimates of decomposition 23 for origin × education subgroups. Immigrants are compared to natives of the same education level. Based on all natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort.

Figure A.14
Assimilation - Comparison with all Natives

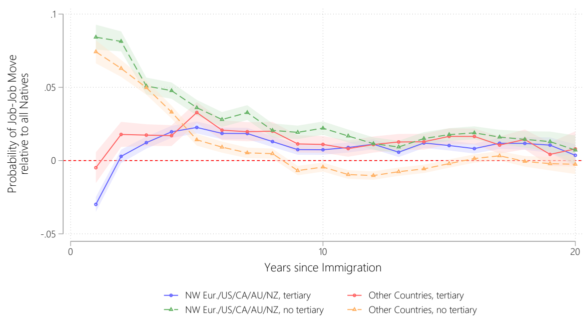
(a) Log Earnings



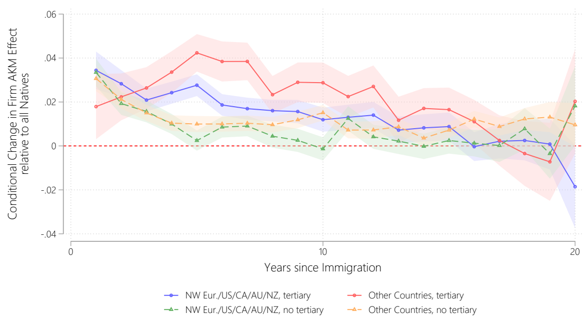
(b) Firm AKM Effect



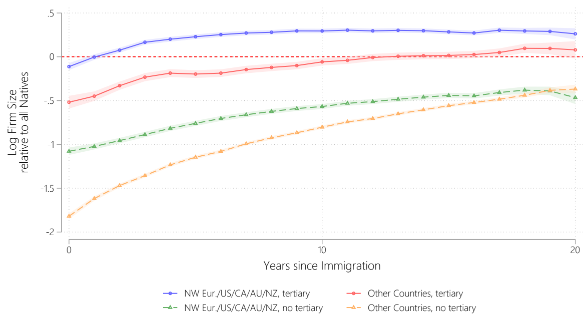
(c) Job-Job Change Probability



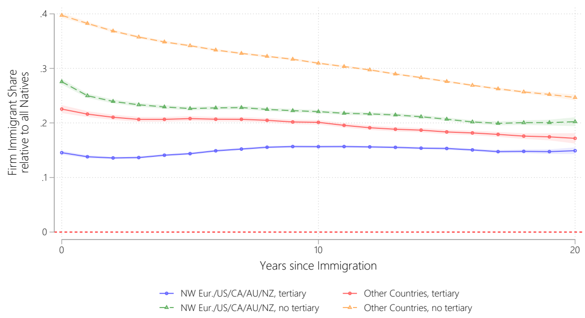
(d) Gain in Firm premium conditional on Job Change



(e) Firm Size

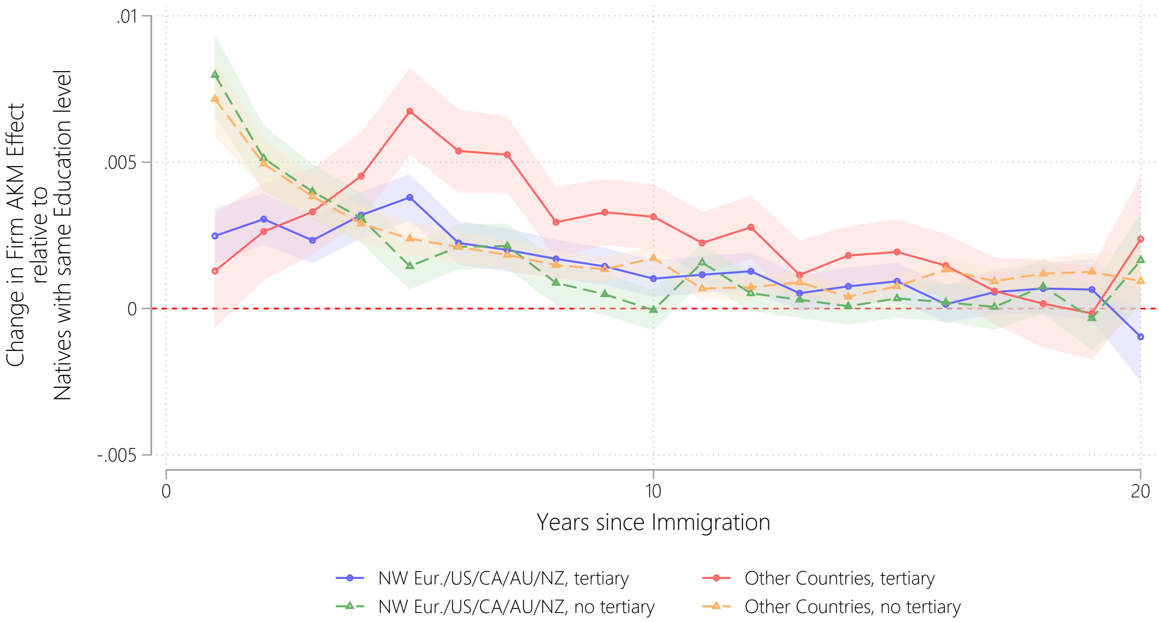


(f) Coworker Immigrant Share



Notes: Point estimates and 95% confidence intervals of $\{\beta_e\}_{e=0}^{20}$ from equations 21, 24, 29, 30, 31 and 32 for origin × education subgroups. Immigrants are compared to all natives. Standard errors are clustered at the individual level. Based on all natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort. Natives' baseline moments:
$\mathbb{E}[\mathbb{1}\{j(igt) \neq j(igt-1)\} \mid N] = 0.11$
$\mathbb{E}[\hat{\psi}^g_{j(igt)} - \hat{\psi}^g_{j(igt-1)} \mid N, switcher] = 0.009$
$\mathbb{E}[\rho_{ij(igt)t} \mid N] = 0.27$

Figure A.15
ASSIMILATION - UNCONDITIONAL GAIN IN FIRM PREMIUM

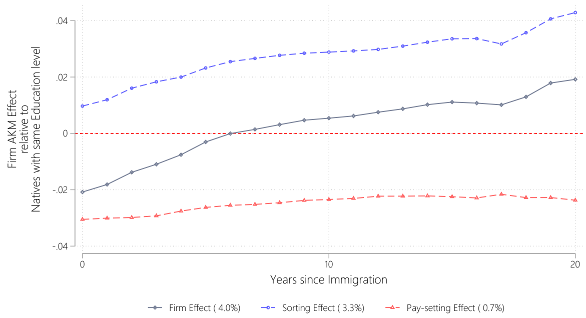


Notes: Point estimates and 95% confidence intervals of $\{\beta_e\}_{e=0}^{20}$ from equation 30 without conditioning on employer changes for origin × education subgroups. Immigrants are compared to natives of the same education level. Standard errors are clustered at the individual level. Based on all natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort.

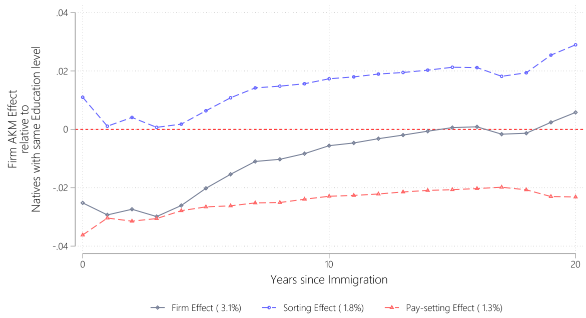
Figure A.16

FIRM EFFECT DECOMPOSITION - STAYERS

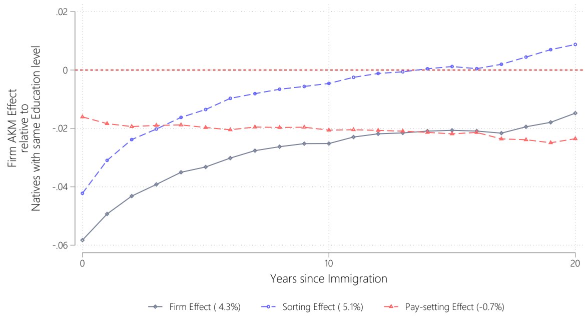
(a) NW EUR./US/CA/AU/NZ
TERTIARY



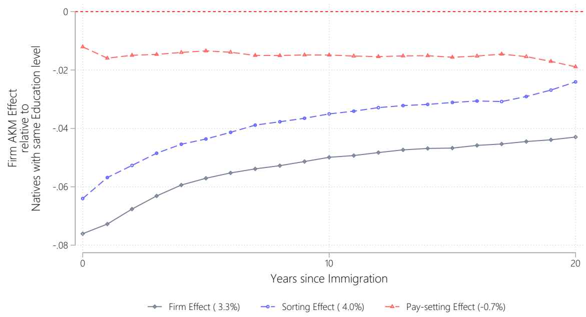
(b) OTHER COUNTRIES
TERTIARY



(c) NW EUR./US/CA/AU/NZ
NO TERTIARY



(d) OTHER COUNTRIES
NO TERTIARY

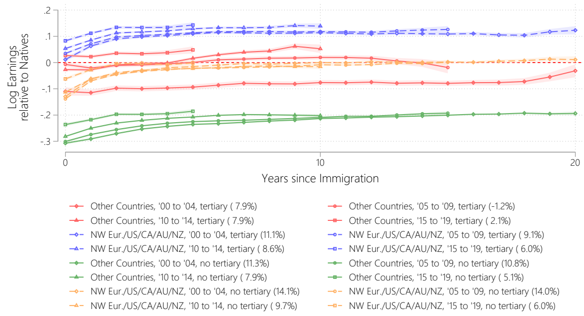


Notes: Point estimates of decomposition 23 for origin × education subgroups. Immigrants are compared to natives of the same education level. Based on all natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort still participating on the labor market in 2020.

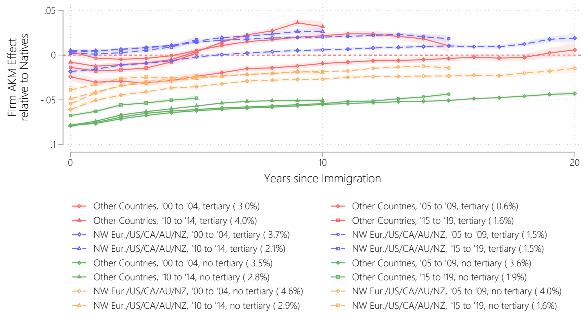
Figure A.17

Assimilation - Cohort × Origin × Education

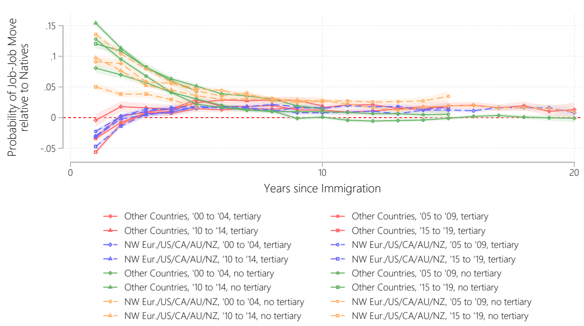
(a) Log Earnings



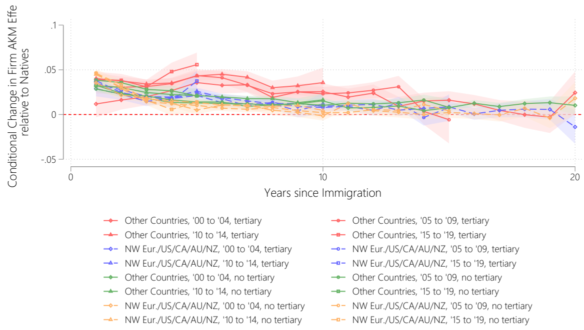
(b) Firm AKM Effect



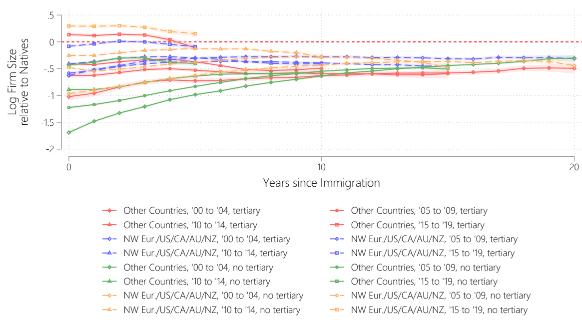
(c) Job-Job Change Probability



(d) Gain in Firm premium conditional on Job Change



(e) Firm Size



(f) Coworker Immigrant Share

Notes: Point estimates and 95% confidence intervals of $\{\beta_e\}_{e=0}^{20}$ from equations 21, 24, 29, 30, 31 and 32 for cohort × origin × education subgroups. Immigrants are compared to natives of the same education level. Standard errors are clustered at the individual level. Based on all natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort. Natives' baseline moments:

$\mathbb{E}[\mathbb{1}\{j(igt) \neq j(igt-1)\} \mid N, tertiary] = 0.11$

$\mathbb{E}[\mathbb{1}\{j(igt) \neq j(igt-1)\} \mid N, non-tertiary] = 0.10$

$\mathbb{E}[\hat{\psi}^g_{j(igt)} - \hat{\psi}^g_{j(igt-1)} \mid N, tertiary, switcher] = 0.012$

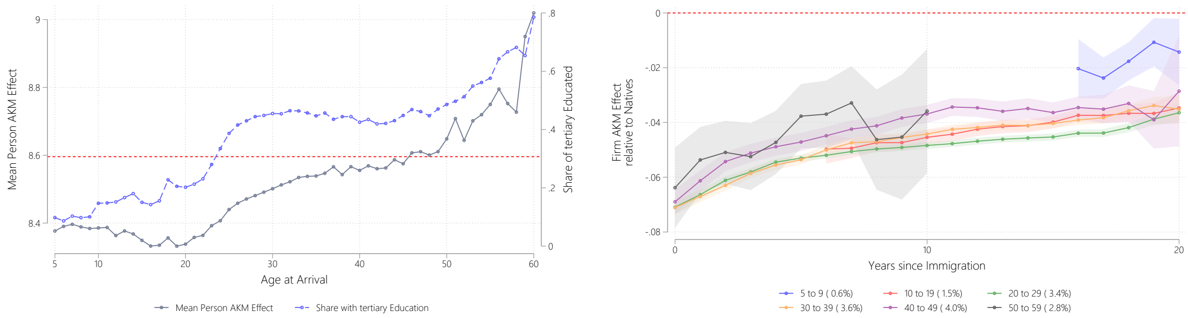
$\mathbb{E}[\hat{\psi}^g_{j(igt)} - \hat{\psi}^g_{j(igt-1)} \mid N, non-tertiary, switcher] = 0.0079$

$\mathbb{E}[\rho_{ij(igt)t} \mid N, tertiary] = 0.26$

$\mathbb{E}[\rho_{ij(igt)t} \mid N, non-tertiary] = 0.28$

Figure A.18
Immigration within critical Development Age

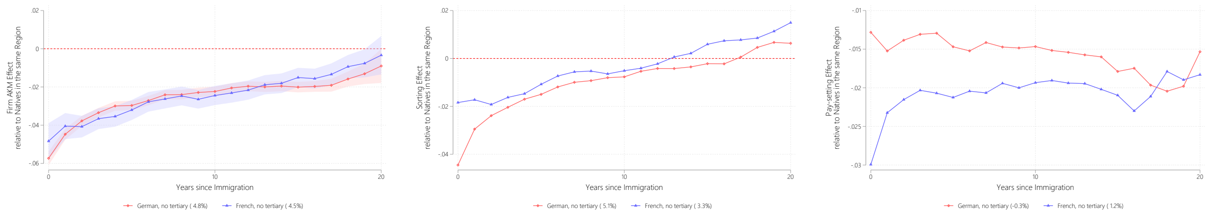
(a) Person AKM Effect by Age at Arrival — **(b)** Firm AKM Effect Low Education



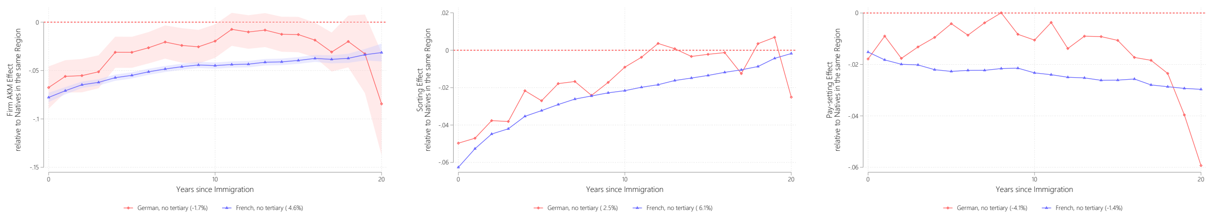
Notes: Panel (a) shows mean immigrant person AKM effects and the share of tertiary-educated immigrants by age at arrival. The dashed red horizontal line indicates the mean natives person AKM effect. Panel (b) shows point estimates and 95% confidence intervals of $\{\beta_e\}_{e=0}^{20}$ from equation 21 for age-at-arrival subgroups. Only non-tertiary-educated immigrants and natives are considered. Immigrants are compared to natives of the same education level. Based on all natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort.

Figure A.19
Assimilation - by Cultural Regions
Non-tertiary Education

(a) German-speaking



(b) French-speaking

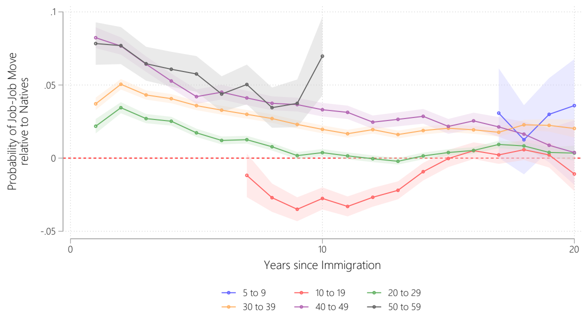


Notes: Point estimates and 95% confidence intervals of $\{\beta_e\}_{e=0}^{20}$ from equation 21 and point estimates of decomposition 23 for non-tertiary-educated German, Austrian and French immigrants in German- and French-speaking labor market regions. Immigrants are compared to non-tertiary-educated natives residing in the defined regions. Standard errors are clustered at the individual level. Based on natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort who reside in the respective German- or French-speaking regions. German-speaking regions are: Bern, Basel, Aarau-Olten, Zurich, Winterthur-Schaffhausen, St. Gallen, Chur, Lucerne. French-speaking regions are: Geneva, Lausanne, Sion, Fribourg, Neuchâtel.

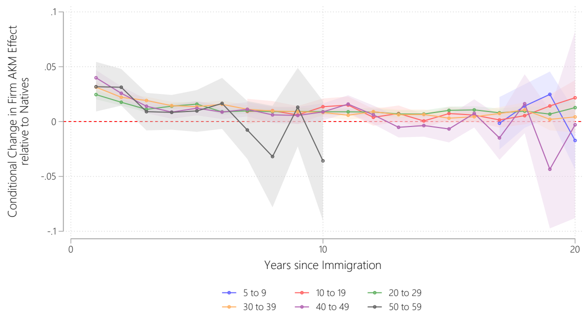
Figure A.20

ADDITIONAL OUTCOMES - AGE AT IMMIGRATION

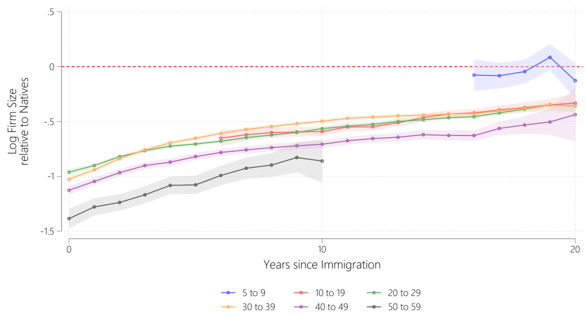
(a) JOB-JOB CHANGE PROBABILITY



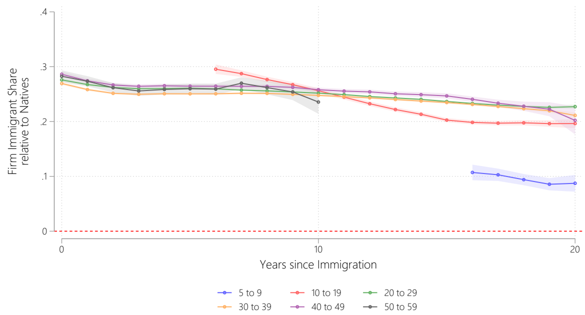
(b) GAIN IN FIRM PREMIUM CONDITIONAL ON JOB CHANGE



(c) FIRM SIZE



(d) COWORKER IMMIGRANT SHARE



Notes: Point estimates and 95% confidence intervals of $\{\beta_e\}_{e=0}^{20}$ from equations 29, 30, 31 and 32 by age-at-arrival subgroups. Immigrants are compared to all natives. Standard errors are clustered at the individual level. Based on all natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort. Natives' baseline moments:

$\mathbb{E}[\mathbb{1}\{j(igt) \neq j(igt-1)\} \mid N] = 0.11$

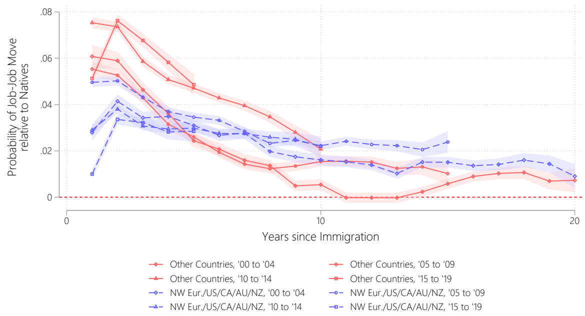
$\mathbb{E}[\hat{\psi}^g_{j(igt)} - \hat{\psi}^g_{j(igt-1)} \mid N, switcher] = 0.009$

$\mathbb{E}[\rho_{ij(igt)t} \mid N] = 0.27$

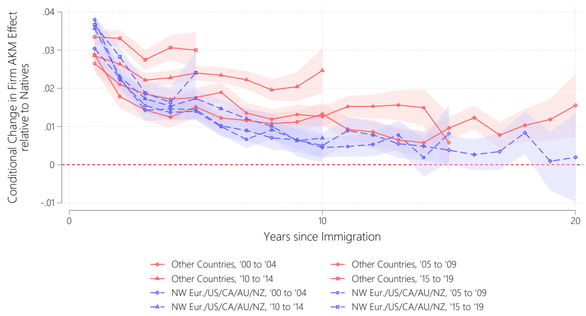
Figure A.21

ADDITIONAL OUTCOMES - COHORT × ORIGIN

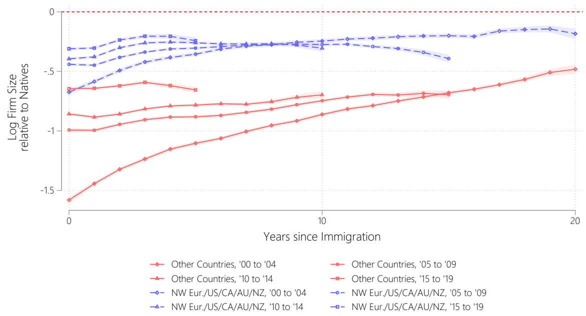
(a) JOB-JOB CHANGE PROBABILITY



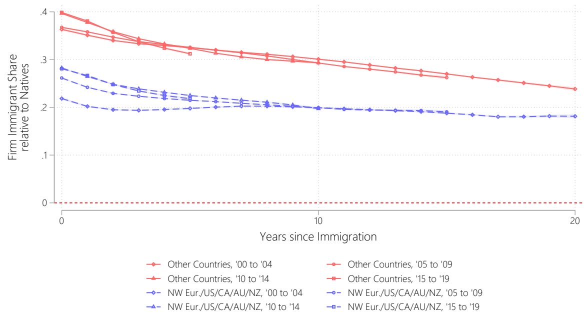
(b) GAIN IN FIRM PREMIUM CONDITIONAL ON JOB CHANGE



(c) FIRM SIZE



(d) COWORKER IMMIGRANT SHARE



Notes: Point estimates and 95% confidence intervals of $\{\beta_e\}_{e=0}^{20}$ from equations 29, 30, 31 and 32 by cohort × origin subgroups. Immigrants are compared to all natives. Standard errors are clustered at the individual level. Based on all natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort. Natives' baseline moments:

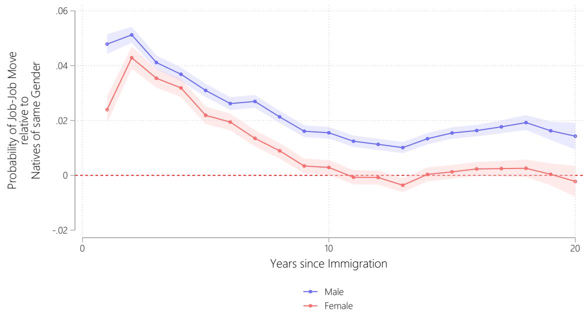
$\mathbb{E}[\mathbb{1}\{j(igt) \neq j(igt-1)\} \mid N] = 0.11$

$\mathbb{E}[\hat{\psi}^g_{j(igt)} - \hat{\psi}^g_{j(igt-1)} \mid N, switcher] = 0.009$

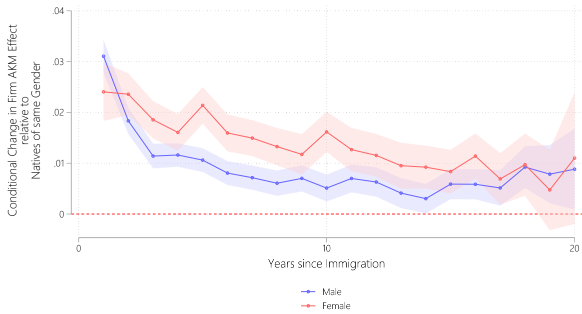
$\mathbb{E}[\rho_{ij(igt)t} \mid N] = 0.27$

Figure A.22
Additional Outcomes - Gender

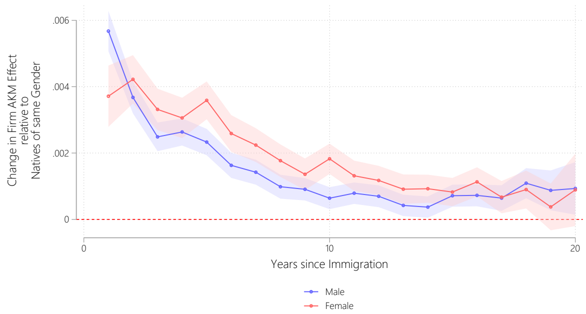
(a) Job-Job Change Probability



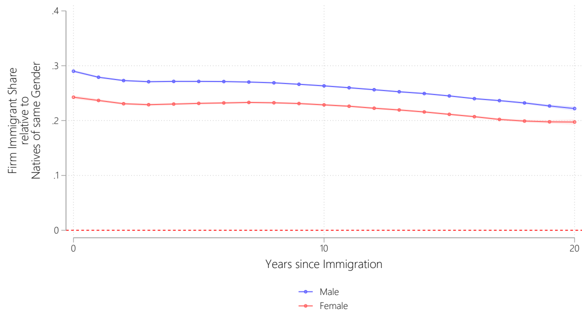
(b) Gain in Firm premium conditional on Job Change



(c) Gain in Firm premium unconditional



(d) Coworker Immigrant Share



Notes: Point estimates and 95% confidence intervals of $\{\beta_e\}_{e=0}^{20}$ from equations 29, 30, 31 and 32 by gender subgroups. Immigrants are compared to natives of the same gender. Standard errors are clustered at the individual level. Based on all natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort. Natives' baseline moments:
$\mathbb{E}[\mathbb{1}\{j(igt) \neq j(igt-1)\} \mid N, male] = 0.11$
$\mathbb{E}[\mathbb{1}\{j(igt) \neq j(igt-1)\} \mid N, female] = 0.11$
$\mathbb{E}[\hat{\psi}^g_{j(igt)} - \hat{\psi}^g_{j(igt-1)} \mid N, male, switcher] = 0.011$
$\mathbb{E}[\hat{\psi}^g_{j(igt)} - \hat{\psi}^g_{j(igt-1)} \mid N, female, switcher] = 0.006$
$\mathbb{E}[\rho_{ij(igt)t} \mid N, male] = 0.27$
$\mathbb{E}[\rho_{ij(igt)t} \mid N, female] = 0.27$

A.2 Tables

Table A.1
Region Fixed Effect Estimates
from AKM Models

	Estimates	
Region	Natives	Immigrants
Geneva	0	0
Lausanne	-0.0177	0.00578
Sion	-0.0379	0.00115
Fribourg	-0.0334	0.000634
Neuchâtel	-0.0344	0.00243
Biel/Bienne	-0.0406	0.00498
Bern	-0.0345	0.00887
Basel	-0.0320	-0.00113
Aarau-Olten	-0.0387	0.0144
Zurich	-0.00779	0.0333
Winterthur-Schaffhausen	-0.0316	0.0141
St. Gallen	-0.0492	-0.00266
Chur	-0.0478	0.00517
Lucerne	-0.0366	0.0155
Bellinzona	-0.0573	-0.00488
Lugano	-0.0462	-0.00808
Missing	-0.0306	-0.0199

Notes: Estimates of labor market region fixed effects obtained from the AKM model. Based on the largest connected set for each origin group.

Table A.2
CRE and BLM Bias Correction

	Samples	
	Natives (1)	Immigrants (2)
Panel A: CRE		
Firm effect	6.0%	6.8%
Covariance of person and firm effects	12.6%	21.4%
Panel B: BLM		
Firm effect	4.5%	7.0%
Covariance of person and firm effects	12.0%	20.8%

Notes: Panel A shows results of the CRE model proposed by Bonhomme et al. (2023) with $k = 10$ groups. The log earnings are first residualized using the baselines covariates, afterwards the model is fit, see appendix B.2 of Bonhomme et al. (2023) for more information. Panel B shows results for the BLM model proposed by Bonhomme, Lamadon, and Manresa (2019) with $k = 10$ and restricted to no interaction between firm groups and worker effects.

Table A.3
AKM Decomposition - No regional Controls

	Samples	
	Natives (1)	Immigrants (2)
Largest Connected Set		
Explained log earnings variance		
Person effect	65.2%	60.4%
Firm effect	11.6%	11.8%
Covariance of person and firm effects	3.2%	11.2%
Covariates and associated covariances	5.2%	5.2%
Residual	14.7%	11.4%

Notes: The table shows results for the largest connected set for every origin group. The AKM regression does not use controls for the 16 labor market areas.

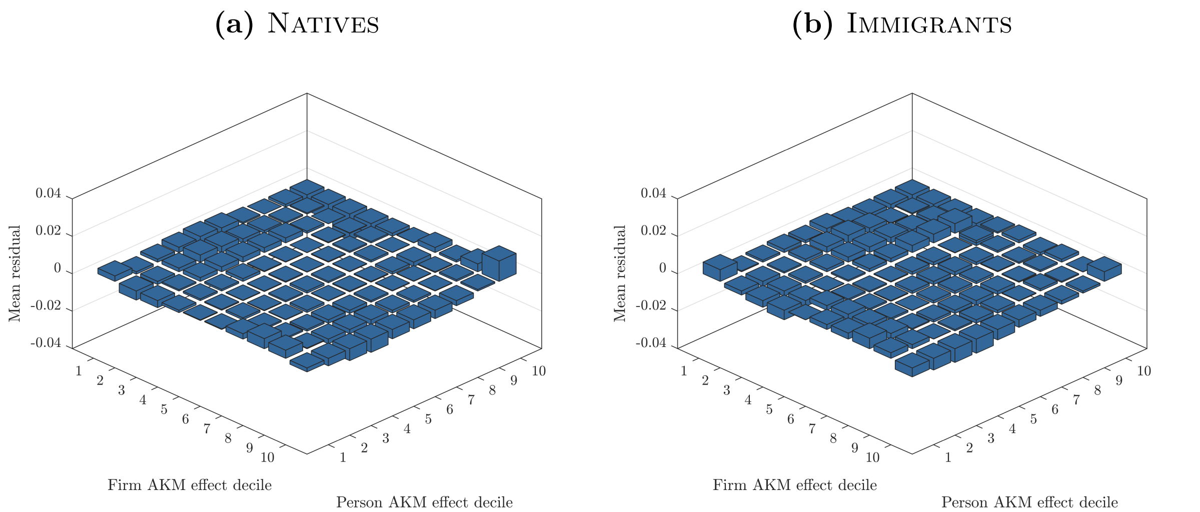
Table A.4
AKM Decomposition - Full Sample

	Natives and Immigrants
	(1)
Panel A: Largest Connected Set	
Mean of log earnings	8.770
Standard deviation of log earnings	0.499
Number of person-year observations	58,027,778
Number of persons	5,911,930
Number of movers	3,306,535
Number of firms	937,325
AKM	
Person effect	64.6%
Firm effect	10.9%
Covariance of person and firm effects	4.3%
Covariates and associated covariances	5.3%
Residual	14.9%
CRE	
Firm effect	6.4%
Covariance of person and firm effects	16.4%
Panel B: Leave-One-Out-Set	
Coverage of the connected set	97.2%
AKM	
Correlation of person/firm effects	0.210
Person effect	62.7%
Firm effect	8.2%
Covariance of person and firm effects	9.5%
Covariates and associated covariances	6.0%
Residual	13.7%
KSS	
Correlation of person/firm effects	0.239
Person effect	60.4%
Firm effect	7.6%
Covariance of person and firm effects	10.2%
Covariates and associated covariances	4.3%
Residual	17.5%

Notes: The table shows results for the largest connected set for the whole sample, irrespective of immigrant status. Panel A shows results for the largest connected set using 'plug-in' estimates and CRE estimates with $k = 10$ groups. Panel B shows 'plug-in' estimates and KSS estimates applied on the leave-one-out set.

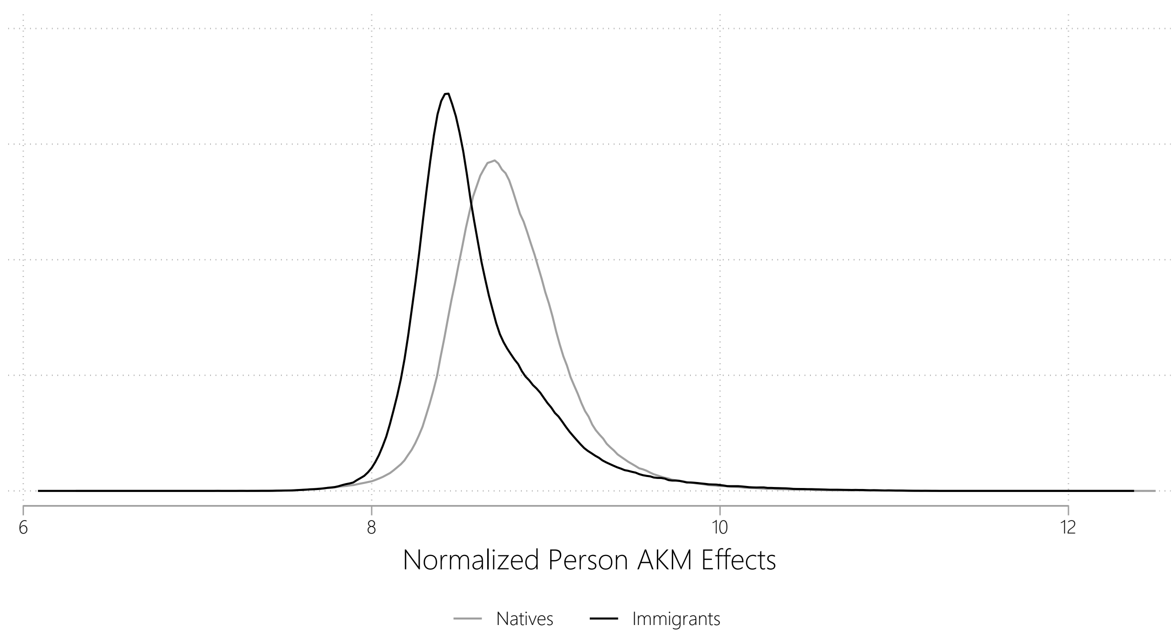
B Male Sample Results

Figure B.1
AKM Residuals by Person and Firm AKM Effect Deciles - Male Sample



Notes: Mean AKM model residuals by person-year weighted firm and person AKM effect deciles. Based on the largest connected set for each origin group.

Figure B.2
Density of Person AKM Effects - Male Sample

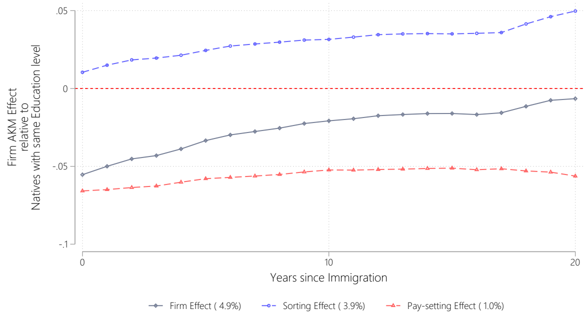


Notes: Estimated person-year weighted density of normalized person AKM effects by origin using epanechnikov kernel and optimal bandwidth. Based on the dual-connected set of the male sample.

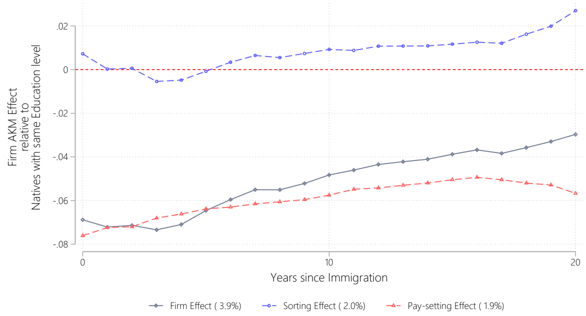
Figure B.3

Firm Effect Change Decomposition - Male Sample

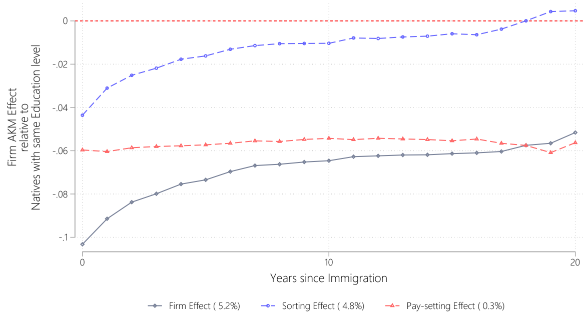
(a) NW Eur./US/CA/AU/NZ
tertiary



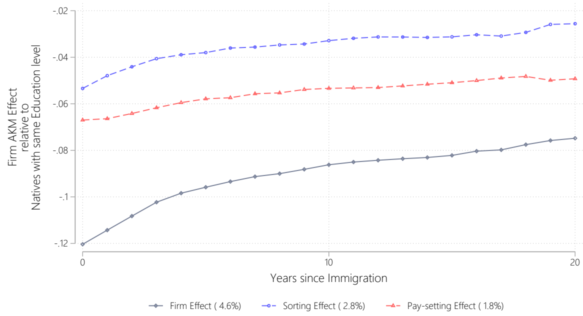
(b) Other Countries
tertiary



(c) NW Eur./US/CA/AU/NZ
no tertiary



(d) Other Countries
no tertiary

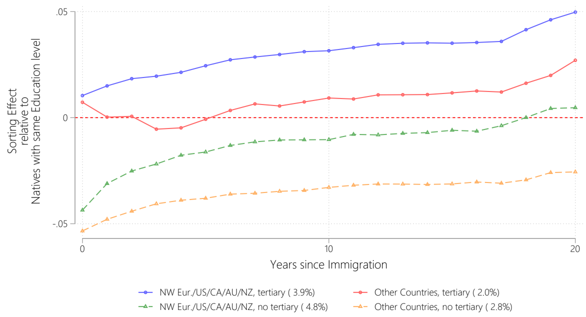


Notes: Estimates of decomposition 23 for origin × education subgroups. Immigrants are compared to natives of the same education level. Based on all natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort of the male sample.

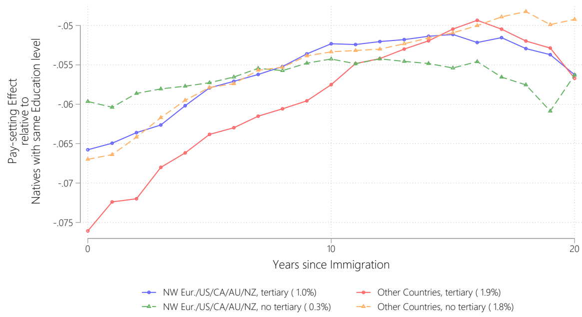
Figure B.4

Sorting and Pay-setting - Male Sample

(a) Sorting



(b) Pay-setting



Notes: Estimates of decomposition 23 for origin × education subgroups. Immigrants are compared to natives of the same education level. Based on all natives in the dual-connected set and immigrants in the 2000–2004 arrival cohort of the male sample.

Table B.1
AKM Decomposition - Male Sample

	Samples	
	Natives (1)	Immigrants (2)
Panel A: Largest Connected Set		
Mean of log earnings	8.940	8.767
Standard deviation of log earnings	0.473	0.511
Observations		
Number of person-year observations	21,419,901	13,271,670
Number of persons	1,788,466	1,541,410
Number of movers	1,073,766	843,976
Number of firms	500,746	341,596
Model Estimates		
Std. dev. of person effects (across person-yr obs.)	0.355	0.389
Std. dev. of firm effects (across person-yr obs.)	0.181	0.183
Std. dev. of covariates (across person-yr obs.)	0.150	0.112
Correlation of person/firm effects	0.022	0.195
Adjusted R-squared	0.844	0.874
RMSE	0.187	0.182
Explained log earnings variance		
Person effect	56.2%	57.9%
Firm effect	14.7%	12.9%
Covariance of person and firm effects	1.3%	10.6%
Covariates and associated covariances	13.9%	7.8%
Residual	13.9%	10.8%
Panel B: Leave-One-Out-Set		
Coverage of the connected set	94.9%	95.9%
AKM		
Correlation of person/firm effects	0.113	0.294
Person effect	56.2%	57.2%
Firm effect	9.6%	10.0%
Covariance of person and firm effects	5.2%	14.1%
Covariates and associated covariances	14.8%	7.9%
Residual	14.1%	10.8%
KSS		
Correlation of person/firm effects	0.153	0.340
Person effect	53.9%	54.8%
Firm effect	8.6%	9.1%
Covariance of person and firm effects	6.6%	15.2%
Covariates and associated covariances	13.0%	6.9%
Residual	17.9%	14.0%
Panel C: Job-Match Effect Model		
Number of job-match effects	4,224,350	3,384,875
RMSE of job-match effect model	0.160	0.160
Adjusted R-squared of job-match effect model	0.885	0.901
Standard deviation of job-match effect	0.097	0.087

Notes: The table is based on the male sample. Panel A shows results of decomposition 18 for the largest connected set for every origin group. Panel B shows the same results for the leave-one-out-set, i.e. the largest connected set with at least two movers in every firm. Additionally Panel B shows bias-corrected estimates using the method outlined in KSS. The job-match effect model controls for a fixed effect for every person-firm match instead of person and firm fixed effects.

Table B.2
AKM Earnings Gap Decomposition - Male Sample

	Transferable Skills			Firms				
	Differece in Log Earnings (1)	Differece in Person Effect (2)	Differece in Log Covariates (3)	Swiss Firm Effect (4)	Immigrant Firm Effect (5)	Differece in Firm Effect (6)	Sorting (7)	Pay-setting (8)
All	0.173	0.185 106.9%	-0.073 -42.2%	0.206	0.144	0.062 35.8%	0.010 5.8%	0.051 29.5%
By Age:								
30 or younger	0.090	0.177 196.7%	-0.160 -177.8%	0.191	0.117	0.074 82.2%	0.018 20.0%	0.056 62.2%
31 to 49	0.178	0.183 102.8%	-0.065 -36.5%	0.210	0.149	0.061 34.3%	0.010 5.6%	0.051 28.7%
50 or older	0.198	0.193 97.5%	-0.051 -25.8%	0.205	0.150	0.055 27.8%	0.007 3.5%	0.048 24.2%
By Education:								
No Tertiary Education	0.182	0.195 107.1%	-0.083 -45.6%	0.202	0.131	0.071 39.0%	0.020 11.0%	0.051 28.0%
Tertiary Education	-0.017	0.035 -205.9%	-0.077 452.9%	0.223	0.198	0.025 -147.1%	-0.030 176.5%	0.055 -323.5%

Notes: Decomposition of immigrant-native earnings gap using equations 19 and 20. Based on the dual-connected set of the male sample.

C Additional Derivations

C.1 Estimation of Assimilation Decomposition

Equation 23 decomposes the estimated firm premium gap by years since migration into components driven by differential sorting and differential pay-setting according to:

$$\begin{aligned}\beta_e =& \mathbb{E}[\psi^N_{j(igt)} \mid g(i) = M, E_{it} = e, X_{it}] - \mathbb{E}[\psi^N_{j(igt)} \mid g(i) = N, X_{it}] \\ &+ \mathbb{E}[\psi^M_{j(igt)} - \psi^N_{j(igt)} \mid g(i) = M, E_{it} = e, X_{it}].\end{aligned} \tag{C.1}$$

This section shows how exactly the components are computed.[C.1]

For reasons of consistency with the overall firm premium gap, notice the above equation conditioning on X_{it}, estimated firm premia are residualized using $\hat{\theta}$ recovered from estimation of equation 21:

$$\begin{aligned}\tilde{\hat{\psi}}^M_{j(igt)} &= \hat{\psi}^M_{j(igt)} - X_{it}\hat{\theta} \\ \tilde{\hat{\psi}}^N_{j(igt)} &= \hat{\psi}^N_{j(igt)} - X_{it}\hat{\theta},\end{aligned} \tag{C.2}$$

i.e. every (i,t)-observation has its immigrant and native firm premium $\hat{\psi}^M_{J(i,t)}$ and $\hat{\psi}^N_{J(i,t)}$ and *one* covariate prediction $X'_{it}\hat{\theta}$ based on the estimated year, age and regional effects. Assuming this, differential pay-setting can be estimated using sample moments as:

$$\begin{aligned}\hat{\mathbb{E}}[\psi^M_{j(igt)} - \psi^N_{j(igt)} \mid g(i) = M, E_{it} = e, X_{it}] &= \frac{\sum_{it} \mathbb{1}\{E_{it} = e, g(i) = M\} \cdot (\tilde{\hat{\psi}}^M_{j(igt)} - \tilde{\hat{\psi}}^N_{j(igt)})}{\sum_{it} \mathbb{1}\{E_{it} = e, g(i) = M\}} \\ &= \frac{\sum_{it} \mathbb{1}\{E_{it} = e, g(i) = M\} \cdot (\hat{\psi}^M_{j(igt)} - \hat{\psi}^N_{j(igt)})}{\sum_{it} \mathbb{1}\{E_{it} = e, g(i) = M\}}.\end{aligned} \tag{C.3}$$

This means, that with definitions given in equation C.2, the residualization does not make a difference for pay-setting.

Differential sorting can be estimated as:

[C.1] Thanks to Jaime Arellano-Bover and Shmuel San for clarifying some unclear points.

$$
\begin{aligned}
&\hat{\mathbb{E}}[\psi^N_{j(igt)} \mid g(i)=M, E_{it}=e, X_{it}] - \hat{\mathbb{E}}[\psi^N_{j(igt)} \mid g(i)=N, X_{it}] \\
&\qquad = \frac{\sum_{it} \mathbb{1}\{E_{it}=e, g(i)=M\} \cdot \tilde{\hat{\psi}}^N_{j(igt)}}{\sum_{it} \mathbb{1}\{E_{it}=e, g(i)=M\}} - \frac{\sum_{it} \mathbb{1}\{g(i)=N\} \cdot \tilde{\hat{\psi}}^N_{j(igt)}}{\sum_{it} \mathbb{1}\{g(i)=N\}} \\
&\qquad = \frac{\sum_{it} \mathbb{1}\{E_{it}=e, g(i)=M\} \cdot \hat{\psi}^N_{j(igt)}}{\sum_{it} \mathbb{1}\{E_{it}=e, g(i)=M\}} - \frac{\sum_{it} \mathbb{1}\{g(i)=N\} \cdot \hat{\psi}^N_{j(igt)}}{\sum_{it} \mathbb{1}\{g(i)=N\}} \\
&\qquad\quad - \underbrace{\left[\frac{\sum_{it} \mathbb{1}\{E_{it}=e, g(i)=M\} \cdot X_{it}\hat{\theta}}{\sum_{it} \mathbb{1}\{E_{it}=e, g(i)=M\}} - \frac{\sum_{it} \mathbb{1}\{g(i)=N\} \cdot X_{it}\hat{\theta}}{\sum_{it} \mathbb{1}\{g(i)=N\}}\right]}_{\text{composition}}.
\end{aligned} \tag{C.4}
$$

This means that sorting depends on the expected/predicted firm premium (based on age, year and regional effects) of immigrants active in year since arrival e relative to all natives. Intuitively, a more favorable composition of immigrants in terms of, for example, age relative to natives, decreases the estimated sorting effect, because some of the 'better' sorting can be explained by compositional effects.

For reasons of computational simplicity, it is possible to compute either one of the effects and obtain the other effect as the residual.

C.2 Firm Entry and Exit Decomposition

Change in Effects. Let's define the growth in expected earnings between year s and year $t > s$ for group g as:

$$\begin{aligned}\Delta_{gts} \equiv \mathbb{E}[\ln y_{igt}] - \mathbb{E}[\ln y_{igs}] = \quad & \mathbb{E}[\alpha_{ig} \mid \mathbb{1}_{igt} = 1] - \mathbb{E}[\alpha_{ig} \mid \mathbb{1}_{igs} = 1] \\ & + (\bar{X}_{gt} - \bar{X}_{gs})\beta_g \\ & + \sum_j \psi_j^g (\pi_{gjt} - \pi_{gjs}).\end{aligned} \tag{C.5}$$

In a further step it is possible to write the excess growth of immigrant earnings relative to natives as:

$$\begin{aligned}\Delta_{Nts} - \Delta_{Mts} = \quad & \mathbb{E}[\alpha_{iN} \mid \mathbb{1}_{iNt} = 1] - \mathbb{E}[\alpha_{iN} \mid \mathbb{1}_{iNs} = 1] \\ & - (\mathbb{E}[\alpha_{iM} \mid \mathbb{1}_{iMt} = 1] - \mathbb{E}[\alpha_{iM} \mid \mathbb{1}_{iMs} = 1]) \\ & + (\bar{X}_{Nt} - \bar{X}_{Ns})\beta_N - (\bar{X}_{Mt} - \bar{X}_{Ms})\beta_M \\ & + \underbrace{\sum_j \psi_j^N \Delta\pi_{Njts} - \sum_j \psi_j^M \Delta\pi_{Mjts}}_{\Delta \text{ firm effect}},\end{aligned} \tag{C.6}$$

where $\Delta\pi_{gjts} = \pi_{gjt} - \pi_{gjs}$. $\Delta_{Nts} - \Delta_{Mts}$, i.e. the mean difference in earnings growth rates between natives and immigrants can be interpreted as a measure of earnings assimilation. It is possible to track changes in sorting and pay-setting effects by applying a similar decomposition as in equation 20 to the change in firm effect in equation C.6:

$$\sum_j \psi_j^N \Delta\pi_{Njts} - \sum_j \psi_j^M \Delta\pi_{Mjts} = \underbrace{\sum_j \psi_j^N (\Delta\pi_{Njts} - \Delta\pi_{Mjts})}_{\Delta \text{ sorting effect}} + \underbrace{\sum_j (\psi_j^N - \psi_j^M)\Delta\pi_{Mjts}}_{\Delta \text{ pay-setting effect}}. \tag{C.7}$$

Defined in this way, a *decrease* of the assimilation measure corresponds to a *more favorable* situation for immigrants. This is important to keep in mind.

Firm Entry and Exit. Using the decomposition Haltiwanger (1997) originally used to decompose total factor productivity it is possible to decompose the change in any weighted average $\bar{\mu}_t = \sum_i m_{it}\mu_{it}$, where $\sum_i m_{it} = 1$ into components driven by the change in the measure on the firm level $\Delta\mu_{it}$ (within component), the change in weights Δm_{it} (between component), the covariance component of these two $\Delta m_{it}\Delta\mu_{it}$ and components driven by entry of new and exit of existing firms. When applying the decomposition to the case of time-invariant firm premia, the within and covariance components are naturally zero.

In the following I will show how to derive the decomposition for the change in the sorting effect.
Let us begin with the definition of the change in sorting effect from equation C.7 and set $s = t - 1$:

$$\begin{aligned}\Delta\bar{\psi}_t^{\text{sorting}} &= \sum_j \psi_j^N(\Delta\pi_{Njts} - \Delta\pi_{Mjts}) \\ &= \sum_j \psi_j^N(\pi_{Njt} - \pi_{Njt-1} - \pi_{Mjt} + \pi_{Mjt-1}),\end{aligned} \tag{C.8}$$

where $\mathcal{J}_t$ is the set of firms active in period t, $\bar{\psi}_{t-1}^{\text{sorting}} \equiv \sum_j \psi_j^N(\pi_{Njt} - \pi_{Mjt})$ and $\Delta\bar{\psi}_t^{\text{sorting}} \equiv \bar{\psi}_t^{\text{sorting}} - \bar{\psi}_{t-1}^{\text{sorting}}$. By adding and subtracting $\bar{\psi}_{t-1}^{\text{sorting}}$ and using the fact that $\sum_{j\in\mathcal{J}_t} \pi_{gjt} = 1 \ \forall\ t, g \in \{N, M\}$:

$$\begin{aligned}\Delta\bar{\psi}_t^{\text{sorting}} &= \sum_{j\in\mathcal{J}_t} \psi_j^N(\pi_{Njt} - \pi_{Mjt}) - \bar{\psi}_{t-1}^{\text{sorting}} + \bar{\psi}_{t-1}^{\text{sorting}} \\ &\quad - \sum_{j\in\mathcal{J}_{t-1}} \psi_j^N(\pi_{Njt-1} - \pi_{Mjt-1}) - \bar{\psi}_{t-1}^{\text{sorting}} + \bar{\psi}_{t-1}^{\text{sorting}} \\ &= \sum_{j\in\mathcal{J}_t} (\psi_j^N - \bar{\psi}_{t-1}^{\text{sorting}})(\pi_{Njt} - \pi_{Mjt}) \\ &\quad - \sum_{j\in\mathcal{J}_{t-1}} (\psi_j^N - \bar{\psi}_{t-1}^{\text{sorting}})(\pi_{Njt-1} - \pi_{Mjt-1}).\end{aligned} \tag{C.9}$$

Considering those firms in both, $\mathcal{J}_t$ and $\mathcal{J}_{t-1}$, using a separate sum operator one gets:

$$\begin{aligned}\Delta\bar{\psi}_t^{\text{sorting}} &= \sum_{j\in\mathcal{J}_{t-1}\cap\mathcal{J}_t} (\psi_j^N - \bar{\psi}_{t-1}^{\text{sorting}})(\Delta\pi_{Njt} - \Delta\pi_{Mjt}) \\ &\quad + \sum_{j\notin\mathcal{J}_{t-1}\wedge j\in\mathcal{J}_t} (\psi_j^N - \bar{\psi}_{t-1}^{\text{sorting}})(\pi_{Njt} - \pi_{Mjt}) \\ &\quad - \sum_{j\in\mathcal{J}_{t-1}\wedge j\notin\mathcal{J}_t} (\psi_j^N - \bar{\psi}_{t-1}^{\text{sorting}})(\pi_{Njt-1} - \pi_{Mjt-1}).\end{aligned} \tag{C.10}$$

By defining $\tilde{\psi}_{it}^N \equiv \psi_j^N - \bar{\psi}_{t-1}^{\text{sorting}}$ it is possible to write the decomposition as shown in equation 28. The decomposition of the change in pay-setting effect over time can be derived analogously.

D Additional Assimilation Characteristics

Gibbs, Ierulli, and Milgrom (2003) show using Swedish data that occupational changes and promotions contribute to earnings growth. Fox (2009), also using Swedish data, records that earnings increase with job responsibility and job responsibility amplifies earnings-caused differences by firm sizes.

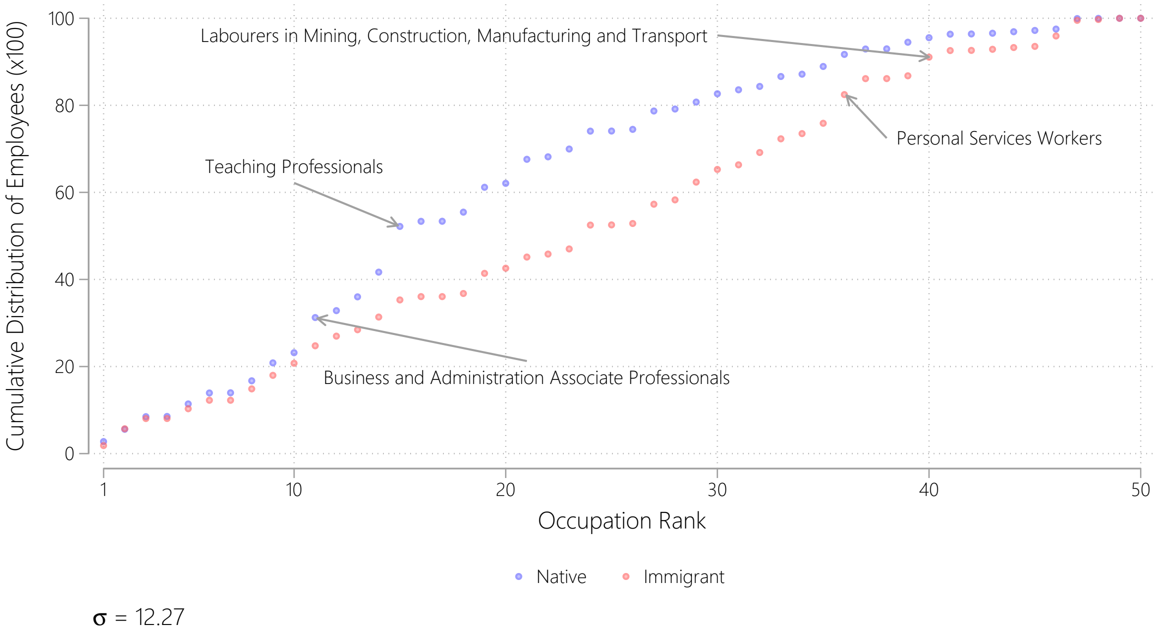
ESS and SE surveys include data on the current occupation, additionally the ESS includes data on the managerial responsibility of the worker measured on a scale of 1, 'senior manager', to 5, 'no responsibility'. I construct an occupation rank variable by residualizing log earnings by age and year effects and computing average log earnings by 2-digit occupations, afterwards a rank is assigned, with 1 being the occupation with the highest residualized earnings.

The cumulative distribution of occupation ranks is shown in figure D.1 separately for natives and immigrants. It is evident that distributions of natives and immigrants are quite similar for the highest-paying occupations. The distributions slowly begin diverging after occupational rank 5 and keeps diverging further until approximately occupational rank 25. Natives are especially overrepresented in the high-ranking occupations 'teaching professionals' and 'business and administration associate professionals'. Immigrants are overrepresented in the lower half of occupations, especially so as 'labourers in mining, construction, manufacturing and transport' and 'personal service workers'.

Figure D.2 shows shares of natives and immigrants across the five managerial responsibility ranks. Natives are overrepresented in all managerial ranks.

Figure D.1

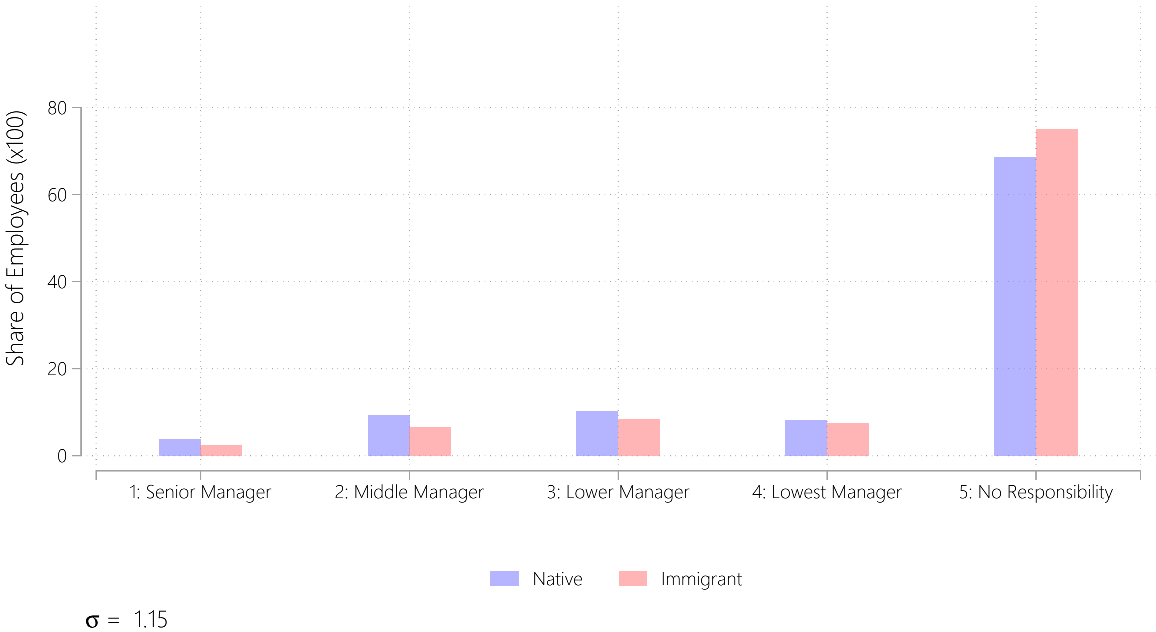
Cumulative Distribution of Occupation Rank



Notes: The figure shows the cumulative distribution of employees by occupation rank separately for natives and immigrants. The construction is based on the male sample to reduce the influence of part-time work. Certain occupations with a higher relative share of either group are labeled. Current occupations are obtained from the SE and ESS surveys, consequently only a fraction of person-year observations are used. The pooled standard deviation is given below the figure. Based on all natives in the dual-connected set and immigrants in the 2008–2012 arrival cohort.

Figure D.2

Distribution of Managerial Responsibility Rank



Notes: The figure shows the distribution of employees by managerial position separately for natives and immigrants. Current managerial positions are obtained from the ESS survey, consequently only a fraction of person-year observations are used. The pooled standard deviation is given below the figure. Based on all natives in the dual-connected set and immigrants in the 2008–2012 arrival cohort.

I estimate the following equations to measure climbing in occupational ranks and managerial positions of immigrants relative to natives:

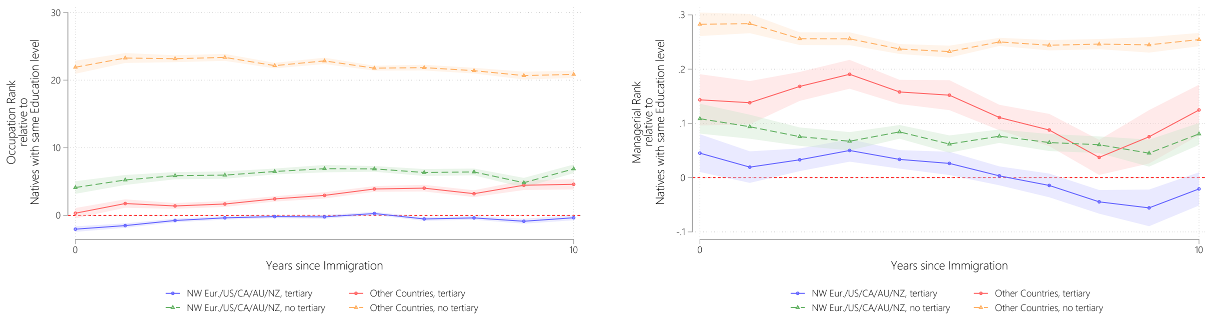
$$occrank_{igj(igt)t} = \mathbb{1}\{g(i) = M\} \times \left[\sum_{e=0}^{10} \beta_e \mathbb{1}\{E_{it} = e\} \right] + X_{it}\theta + \varepsilon_{it}, \tag{D.1}$$

$$position_{igj(igt)t} = \mathbb{1}\{g(i) = M\} \times \left[\sum_{e=0}^{10} \beta_e \mathbb{1}\{E_{it} = e\} \right] + X_{it}\theta + \varepsilon_{it}, \tag{D.2}$$

where $occrank_{igj(igt)t}$ is the occupational rank and $position_{igj(igt)t}$ the managerial responsibility rank of person i in year t. I estimate the equations separately for origin × education subgroups. Immigrants are compare to natives of the same education level. The ESS is conducted in the years 2012, 2014, 2016 and 2018. The SE is conducted annually since 2010. To measure assimilation from the first year since arrival onwards I use the cohort of immigrants arriving in the time span of 2008–2012 from the dual-connected set.

Figure D.3
ADDITIONAL OUTCOMES - JOB

(a) OCCUPATION RANK **(b)** MANAGERIAL RESPONSIBILITY RANK



Notes: Panel (a) shows point estimates and 95% confidence intervals of $\{\beta_e\}_{e=0}^{10}$ from equation D.1, panel (b) shows point estimates and 95% confidence intervals of $\{\beta_e\}_{e=0}^{10}$ from equation D.2, both for origin × education subgroups. Immigrants are compared to natives of the same education level. Heteroskedasticity-robust standard errors are used.

Figures D.3(a) and D.3(b) show estimates of equations D.1 and D.2 respectively. Tertiary-educated immigrants and immigrants from economically well developed countries have higher occupational ranks at arrival.

The group of tertiary-educated 'NW Eur./US/CA/AU/NZ' immigrants has essentially the same occupational rank as their native peers. Differences between the subgroups are more profound for non-tertiary-educated immigrants. Immigrants' occupational rank does not experience big changes. Most of the subgroups even transition to, on average, worse-paying occupations.

In terms of managerial responsibilities, all subgroups get promoted over time, although the change is relatively small compared to the standard deviation of 1.15. In general, 'NW Eur./US/CA/AU/NZ' have managerial responsibility ranks more similar to natives.

Overall, there is no strong evidence suggesting that immigrants assimilate through climbing the 'job ladder'.

E Data

Worker-Flow Approach. To identify firms over time despite the reassignment of firm identifiers I use firm-to-firm worker flows. Algorithm 1 explains the steps used to identify firm identifier reassignment using pseudocode. For the algorithm to be initiated, the master data has to include at most one observation for every person-year-firm combination.

Algorithm 1 Pseudocode Worker-Flow Approach

for $x \in \{1, 2\}$ **do** ▷ Check if the same firm identifier is continued or reassigned
  for $t \in \{1981, \ldots, 2019\}$ **do**
    `<join data from` t `and` $t + x$ `by person identifier>`
    `<calculate the share of employees present in both` t `and` $t + x$ `for every firm identifier combination>`
    `<assign new common firm identifier if observations have the same old firm identifier and share at least` 20% `of employees>`
    `<update firm identifiers in master data>`
  end for
end for

for $x \in \{0, 1, 2\}$ **do** ▷ Check for firm identifier changes
  for $t \in \{1981, \ldots, 2019\}$ **do**
    `<join data from` t `and` $t + x$ `by person identifier>`
    `<calculate the share of employees present in both` t `and` $t + x$ `for every firm identifier combination>`
    `<assign new common firm identifier if observations share at least` 70% `of employees and have at least` 3 `employees>`
    `<compute transitive closure to identify 'hidden' firm identifier links in the binary relation over the set of firms` $\mathcal{J}$`>`
    `<update firm identifiers in master data>`
  end for
end for

Statutory Declaration

I hereby declare that the thesis with title

Firm Heterogeneity and the Immigrant-Native Earnings Gap

has been composed by myself autonomously and that no means other than those declared were used. In every single case, I have marked parts that were taken out of published or unpublished work, either verbatim or in a paraphrased manner, as such through a quotation.

This thesis has not been handed in or published before in the same or similar form.

Zurich, March 27, 2024