International Review of Economics (2026) 73:45
https://doi.org/10.1007/s12232-026-00550-0

RESEARCH ARTICLE

Analysing occupational migrant segregation in the European Union using a sample selection framework

Pietro Giorgio Lovaglio[1]

Received: 18 June 2025 / Accepted: 8 September 2026
© The Author(s) 2026

Abstract

The European labour market reveals significant disparities between migrant and national workers in terms of employment rates, job quality, and occupational outcomes. This paper proposes a robust approach to analysing migrant segregation and differences in employment outcomes by correcting for selection bias within worker groups, thereby addressing both observed and unobserved factors related to work attitudes among migrants and natives. It also constructs counterfactual distributions, providing reliable insights into working condition disparities. Empirical evidence from the 2024 Labour Force Survey for Italy, France, and Germany shows that national workers generally benefit from better working conditions, high-skilled occupations, and more attractive sectors. In contrast, migrant workers are disproportionately concentrated in low-skilled occupations, including elementary roles, medium-skilled industrial positions, and accommodation and food services, with a notable presence in the household employment sector in Italy. Even after accounting for individual characteristics and using a counterfactual approach, migrants remain overrepresented in low-skilled and service-oriented positions and underrepresented in managerial and professional occupations, particularly in Italy. These findings highlight the persistent cultural and institutional barriers facing migrants, emphasising the need for comprehensive integration policies and labour market reforms.

Keywords Occupational segregation · Migrants workers · Generalized joint regression models · Counterfactualanalysis

JEL Classification C30 · J21 · J15

✉ Pietro Giorgio Lovaglio
piergiorgio.lovaglio@unimib.it

[1] University of Milano-Bicocca, Milan, Italy

Published online: 03 November 2026

1 Introduction

1.1 Labour migration is a key aspect of migration policies

According to the World Migration Report 2024 (McAuliffe and Oucho 2024), there were 169 million international migrant workers globally in 2022, comprising 4.9% of the global labour force. The majority (approximately 67%) were concentrated in high-income countries.

For governments and institutions developing labour market policies targeting migrant populations, it is essential to analyse differences in employment status—including (un)employment rates—and disparities in job opportunities between foreign and national workers.

Prior research has shown that inequalities related to migration status persist across much of the globe (Adsera and Chiswick 2007; Reyneri and Fullin 2011; Biletta et al. 2019; Grubanov-Boskovic et al. 2020). The resulting labour disadvantages, such as occupational segregation, mean that migrants tend to occupy positions at the bottom of the occupational ladder and earn lower wages than their native counterparts (Kaufman 2010; Ballarino and Panichella 2018). Moreover, foreign individuals often experience lower employment rates and labour force participation compared to national populations.

In 2023 (Eurostat 2023a, 2023b), 80.5% of nationals aged 20–64 living in the European Union (EU) were part of the labour force—a higher participation rate than for non-EU citizens (71.8%). This gap is also reflected in the employment (76.2% of nationals vs. 63.0% for non-EU citizens) and unemployment rates (5.4% for national vs. 12.2%, for non-EU citizens) for the same age group.

Research on occupational disparities in European labour markets underscores the interplay of gender, migration, and employment status, all of which shape wage inequality and differences in job quality (Palencia-Esteban and del Río, 2024; Grubanov-Boskovic et al. 2020). Palencia-Esteban (2022) finds that migrant women participate less in the labour market, as they shoulder a disproportionate share of household responsibilities. This, in turn, exacerbates the disadvantages they face relative to native women in terms of both wages and occupations.

Moreover, the impact of immigration on employment varies across European regions, influenced by regional labour market characteristics and demographic composition (Özgüzel and Edo 2023).

This divergent pattern of employment rates between natives and migrants clearly reflects differing levels of employability between the two groups, which cannot be ignored when analysing labour market gaps by ethnicity—such as occupational or sectoral segregation and wage disparities, as well documented by recent studies (Migration Observatory 2023; Broberg et al. 2024).

While institutions have tended to provide general comparisons among foreigners and nationals using such demographics as country of birth, sex, age class, and education, only a handful have compared labour market characteristics (often at aggregate levels), such as classical labour force participation rate or (un)employment rates (Eurostat 2022).

However, when comparing the labour market outcomes of national and migrant workers, simply measuring the difference in these outcomes is insufficient. This approach mainly reflects demographic composition rather than actual labour market dynamics, as it fails to account for differences in individuals' employment propensity.

Employment decisions (and, earlier, labour market participation decisions) are influenced not only by or observed factors such as education, field of study, but also by unobserved factors such as motivation, ambition, health, family obligations. Ignoring selection mechanisms and unobserved factors can lead to biased estimates.

This introduces self-selection bias, as the individuals being compared (employed workers) are not randomly drawn from the general population and workforce composition of migrants and natives systematically differs from their respective broader populations (Glitz 2014; Bertoli et al. 2016).

If self-selection is ignored, comparisons of employment rates and labour market characteristics may produce misleading conclusions, as the observed differences may stem from variations in who chooses—or is able—to work rather than from true labour market disparities (Roy 1951).

For example, since migration often occurs early in life, migrants tend to represent a larger share of the working-age population compared with natives (d'Aiglepierre et al. 2020). Moreover, in family migration contexts, men often migrate first, with other family members following later (Poeschel 2020). This sequential migration pattern contributes to male dominance in specific migrant labour sectors.

The primary objective of this paper is to empirically investigate differences in labour market outcomes between native and migrant workers, with a focus on measuring migrant segregation in Italy, France, and Germany. The analysis accounts for the potential non-random selection of individuals into employment and adopts a counterfactual approach to enable more reliable comparisons between the two groups in terms of their distribution across sectors, occupations, and other labour market outcomes.

The analysis leverages the extensive EU Labour Force Survey (LFS) microdata from the year 2024 (data refer to 2023), the most comprehensive source for assessing labour market trends across European countries, providing valuable insights into employment patterns, population and workforce characteristics (Eurostat 2025).

In this respect, another key contribution of the present paper lies in the use of a comprehensive set of individual- and regional-level covariates—typically omitted in empirical analyses—that help capture the factors shaping differences between migrant and native workers (in terms of employment propensity and labour market outcomes), while accounting for the broader regional context. As also noted by Broberg et al. (2024), the only exception—there is a lack of quantitative research examining how structural and regional labour market characteristics explain employment differences between migrants and nationals in Europe.

Methodologically, the analysis addresses (non-random sample) selection bias by applying a bivariate sample selection model (SSM), ensuring a more accurate assessment of the factors shaping expected employment differences between migrants and national workers.

Specifically, we apply a Generalized Joint Regression Model (GJRM, Smith 2003; Winkelmann 2011; Marra and Radice 2013), which allows for the simultaneous esti-

mation of selection processes (individuals in the labour market) and if the worker is a migrant or a national.

The model consists of two jointly estimated equations: the first (selection equation) assesses the probability of being employed across the entire sample of individuals surveyed in the LFS, while the second estimates the likelihood of being a foreign, rather than a native, worker, conditional on employment status from the first equation.

This is the main contribution of the present study, given that many empirical studies on migrant–national differences have failed to fully address proper selection mechanisms, as highlighted by Pecoraro and Wanner (2019), Brücker et al. (2016), Tani (2017), Glitz (2014), Chiswick (1978), Chiswick (1986), Dustmann and Glitz (2011), Borjas (2014).

As a by-product, a second aim of this paper is to provide a robust assessment of migrant–native differences and patterns of migrant segregation. Segregation is typically measured using the Duncan dissimilarity index, which summarizes differences between the shares of migrant and native workers—relative to the total number in each group—across categories of labour market outcomes.

Our approach extends this framework by not only correcting for selection bias, but also by evaluating a counterfactual scenario, within a potential outcomes' framework (Rubin 2005), in which migrants and natives share identical demographic characteristics.

Specifically, we compare observed segregation in occupations, sectors, and contract types with the segregation predicted by a sample selection model (controlling for the socio-demographic and educational characteristics), as well as with the segregation predicted by combining the sample selection model with the counterfactual approach.

By comparing observed segregation with that derived from actual–counterfactual gaps, we can accurately assess whether migrants are disproportionately concentrated in specific sectors, occupations, or contract types, providing a clearer picture of systematic segregation.

The remainder of the paper is structured as follows. Section 2 reviews the relevant literature on migrant–native disparities in the labour market. Section 3 describes the data sources and key variables used in the analysis. Section 4 outlines the methodological framework, while Sect. 5 presents the empirical findings. Finally, Sect. 6 discusses the policy implications of the results and provides concluding remarks.

2 Background

The European labour market exhibits significant disparities between migrant and native workers regarding employment rates, job quality, wages, and occupational mismatches (Biletta et al. 2019; De la Rica et al. 2015). Foreign workers, particularly non-EU citizens and women, face higher risks of unemployment and precarious employment, resulting in their overrepresentation in low-status, low-skilled occupations that require minimal formal education (Eurostat, 2023c; OECD/European Commission, 2023; Sparreboom and Tarvid 2017). This concentration varies across

countries like Italy, Spain, and Germany (Palencia-Esteban and del Río, 2024; Reyneri and Fullin 2011; Ballarino and Panichella 2018) and is often driven by labour demand in elementary roles and professional services, directing migrants towards physically demanding sectors with chronic shortages, including health, logistics, agriculture, and food services (Fernández-Reino and Sumption 2024; Angrist and Kugler 2003; Kugler and Pica 2006; Fasani and Mazza 2020, 2023; Broberg et al. 2024; Breem and Liebig 2020).

2.1 Causes of ethnic gaps

The literature identifies several mechanisms behind ethnic disparities in labour market outcomes, with strong evidence of migrant disadvantage. Earnings analyses show flatter returns to experience and education acquired abroad, whereby migrants receive lower returns to equivalent skills and are concentrated in lower positions of the earnings distribution (Dustmann et al. 2016). This is reflected in persistent occupational and sectoral segregation, particularly among low-skilled workers, although it may decline over time without disappearing (Glitz 2014). Across Europe, migrants—especially from non-EU countries—display lower employment probabilities and more disadvantaged occupational positions even after controlling for observable characteristics, pointing to structural barriers in labour market access (Dustmann and Frattini 2013).

Complementary explanations emphasize skill mismatch and imperfect transferability of human capital**.** Highly educated migrants often experience overqualification due to limited recognition of foreign credentials (Friedberg 2000; Dustmann et al. 2013). In addition, task-based specialisation suggests that migrants and natives are imperfect substitutes: migrants are more concentrated in manual-intensive jobs, while natives specialize in communication-intensive tasks, reinforcing occupational segregation (Peri and Sparber 2009). These mechanisms jointly motivate the empirical focus on migrant segregation across labour market dimensions.

2.2 Ethnic gap and selection mechanism

Migration and labour market outcomes are shaped by self-selection mechanisms, as emphasized by human capital migration models (Sjaastad 1962; Becker 1964; Chiswick 1999). Migrants differ from natives not only in observable characteristics but also in unobserved traits that influence both migration and employment outcomes (Dustmann and Glitz (2011, 2015); Chiswick 1999; Parey et al. 2017). While most empirical studies focus on wages or sectoral allocation, they often neglect selection into employment, limiting the interpretation of observed gaps (Borjas et al. 2019; Breem and Liebig 2020; Fasani and Mazza 2020, 2023).

Only a limited strand of the literature explicitly accounts for selection in analysing migrant segregation.

Hellerstein and Neumark (2008), Åslund and Skans (2010), and Glitz (2014) measure migrant segregation by comparing observed segregation—using indices such as the Duncan index—with segregation under random or conditional (on covariates) allocation of workers across sectors or municipalities. These approaches account for

differences in observable characteristics between immigrant and native workers, but fail to control for unobserved factors.

Others adopt selection models focusing on migration decisions rather than labour market outcomes (Parey et al. 2017; Gould and Moav 2016; Borjas et al. 2019). However, unobserved factors—such as ability, motivation, or country-specific skills—play a key role in both employment and migrant status, and may generate systematic differences in labour market allocation.

Despite focusing on this specific form of self-selection—while neglecting the selection of individuals into employment—selection on unobservables has long been a central concern in labour and migration economics, providing several insights relevant to our empirical framework.

In standard selection models, the residual in the outcome variable (e.g., in a wage equation) captures individual-specific productivity shocks that arise from unobserved characteristics (Chiswick 1999; Borjas et al. 2019) and interpreting the correlation among residuals of two equations is crucial. The residual may also reflect a particularly good job match between the worker and the current employer/labour market, which generates high returns (Borjas et al. 2019).

Overall, selection can be positive or negative depending on relative returns to skills across countries, but cannot be fully captured by observable characteristics alone (Bertoli et al. 2016; Borjas 1987; Parey et al. 2017). This highlights the importance of modelling selection on unobservables when evaluating migrant segregation, which is central to the empirical strategy adopted in this study.

In our model, the residual from the first-stage selection can be interpreted as an unobserved employability component that influences a worker's pre-selection employment potential within the (host, for migrants) labour market.

The finding of a significant negative residual correlation for example implies that the unobserved characteristics that increase an individual's propensity to work in a given country simultaneously decrease the likelihood that an employed individual is a migrant worker (relative to a native worker). This suggests that high unobserved employability components—broadly defined to include innate ability, ambition, and motivation (Chiswick 1999; Bertoli et al. 2016)—are strongly associated with traits that anchor individuals to the most integrated and successful segments of the host labour market, which are disproportionately characteristic of native workers in the host countries. A high unobserved propensity to work and a negative correlation may reflect high returns to these country-specific skills, which are generally accrued and highly valued by native workers or long-term residents (Gould and Moav 2016).

Considering the broader literature, most approaches either compare migrant and native workers while controlling for random allocation without accounting for unobserved factors, or control for the selection of migrant workers (relative to workers remaining in their country of origin). However, none corrects for the non-random selection of individuals into employment, nor incorporates regional covariates.

The next section outlines our methodological and empirical strategy, beginning with a description of the data and variables used.

3 Data and descriptive statistics

In the EU LFS, the sampling frame consists of individuals who are residents in a country, rather than those who are only physically present at the time of the survey. However, the exact definition of "resident" follows Eurostat's specific criteria: individuals living in private households in the surveyed country or residents who have been (or expect to be) living in the country for at least 12 months (or have already lived there for at least 6 months with the intention of staying a further 12). Hence, provided they meet the residence criteria, the survey includes both nationals and non-nationals**,** including migrants.

The analysis relies on LFS microdata from 2024 for Germany, France, and Italy. The dataset consists of individual-level data, with the sample restricted to interviewed individuals aged 15–64 (LFS variable AGE_GRP).

Two key target variables are defined: labour status (y_1) and migration working status (y_2). The former is derived from the variables WKSTAT (whether the individual worked during the reference week) and STAPRO (employment status in the main job). Specifically, y_1=1 if WKSTAT equals 1 (indicating that the individual worked for pay or profit in the reference week) and STAPRO equals 1, 2, or 3 (as self-employed with employees, self-employed without employees, or an employee, respectively). Otherwise, $y_1 = 0$.

Migration working status (y_2) is defined according to the respondent's country of birth, which represents the standard approach (Bansak et al. 2015). However, when estimating the sample selection model, alternative definitions are also considered—specifically those based on citizenship, as well as migrant subgroups restricted to recent migrants (less than five years of residence) or non-EU migrants. Among workers (y_1=1), individuals classified as migrant workers are foreign-born (y_2 =1); otherwise, they are classified as national workers ($y_2 = 0$).

Using citizenship instead of country of birth to define migrant workers may offer certain advantages. Since foreign-born individuals can acquire citizenship after a legally required period of residence and employment—varying across countries (10 years in Italy for non-EU citizens, 8 years in Germany, and 5 years in France)—this criterion would classify such "new citizens" as nationals. Given that naturalisation generally reflects a stable economic and residential status, treating foreign-born naturalised individuals as nationals can be a reasonable choice when comparing native and migrant workers. However, this approach may also lead to an underestimation of migrant–native disparities, as naturalised individuals often continue to face labour market disadvantages despite their formal citizenship status (Bansak et al. 2015; Borjas et al. 2019).

In the paper, migrants born in EU27, EFTA (Norway, Switzerland, Iceland, Liechtenstein), or born in other European countries (Eastern European countries, Balkans, Russia, Turkey) were summed and indicated as EU_EFTA. Regarding the selection criteria for workers, we excluded those employed in the "Activities of Extraterritorial Organizations and Bodies" (NACE rev.2) sector and in the "Armed Forces Occupations" (ISCO-08, 1D) occupation.

Additionally, to account for regional variability, we employed the Eurostat (2024) database "Regional Indicators." Regional covariates (measured in 2023, except for

wages and GDP) that may affect both outcomes are listed in Table 6 in the Appendix (Sect. A1), which also provides the main literature-based justification. For Italy and France (excluding the Départements d'Outre-Mer), regions are classified at the NUTS2 level, while for Germany, they are classified at the NUTS1 level (Länder).

Obviously, covariates differed between the two equations, as labour (y_1) and migration working status (y_2) capture distinct outcomes.

Labour market characteristics relevant to employment—such as sector, occupation, and similar variables—that primarily distinguish nationals from migrant workers in the second equation cannot be included in the first equation, which is intended to differentiate workers from non-workers regardless of ethnic group. Years of residence cannot be included in the second equation since it automatically identifies migrants.

Instruments (variables that influence the probability of employment without structurally determining an individual's status as a migrant or a native worker) were identified based on the existing literature (see discussion) and validated through their statistical significance.

However, we expect regional factors to primarily influence the employment equation rather than the second equation, and we expect that key education and demographic variables—such as gender, age, education level, and field of study—affect both equations.

For variables lacking clear guidance from the literature, our strategy was to initially include in both equations all variables that could plausibly affect the two outcomes, and then retain them based on their statistical significance.

3.1 Descriptive statistics

In all analyses, including descriptive statistics and model estimations, LFS survey weights were used. The dataset analysed, focusing on individuals aged 15–64 in 2023, represents 37,143,000 observations for Italy (14.1% migrants), 40,399,000 for France (13.1% migrants), and 53,677,000 (23.3% migrants) for Germany. Germany, France, and Italy show similar gender balances (around 50% men in all three countries), but differ in age and education. Germany and Italy have older populations, with higher shares in the 55–59 (12.7% and 12.9%) and 60–64 (11.8% and 11.2%) groups, while France is younger, with a larger share aged 15–19 (10.1%). Educational differences are marked: Italy has more individuals with lower secondary education (32.7%) and fewer with tertiary attainment, whereas tertiary education is more common in France (37.1%) and Germany (29.0%). Migrant presence is highest in Germany (17.7% by citizenship; 23.3% by birthplace), compared to Italy (10.1%; 14.1%) and France (7.8%; 13.1%), with Germany hosting more EU migrants (14.4%) and France more non-European ones (9.2%).

Family structures and migrant profiles also differ. France has the highest share of individuals with children (43.9%), followed by Italy (38.2%) and Germany (35.7%). Across all countries, migrants are more likely to live in families with children—for example, couples with children account for 28.5% of migrants in France (vs 23.4% of natives) and migrant families with young children reach 26% in Germany (vs 19%). In Italy, natives are more often in multi-adult, childless households (38.3%),

Table 1 Job market indicators and characteristics of workers by migration status (Country birth definition)

	DE		FR		IT	
INDICATORS	NAT	MIG	NAT	MIG	NAT	MIG
ACTIVITY RATE	81.7	72.8	74.4	70.0	65.9	71.2
UNEMPLOYMENT RATE	2.6	5.2	6.8	11.6	7.3	10.4
EMPLOYMENT RATE	79.6	69.0	69.3	61.9	61.1	63.8
MALE EMPL RATE	77.3	81.5	69.5	71.0	78.4	69.2
FEMALE EMPL RATE	60.5	77.6	55.5	67.7	51.1	52.8
WORKERS						
N (Thousands)	32,768	8,623	24,352	3,266	19,498	3,337
Individual variables (column %)	%	%	%	%	%	%
AGE_GRP						
Y15-19	3.1	1.2	2.4	0.8	0.6	0.3
Y20-24	8.4	5.1	8.2	4.5	4.7	3.8
Y25-29	9.9	9.6	10.4	9.4	8.1	8.1
Y30-34	11.0	13.1	11.6	11.2	10.0	11.3
Y35-39	10.4	14.8	12.1	13.2	10.4	14.0
Y40-44	10.6	13.9	12.4	14.7	12.0	15.9
Y45-49	9.3	12.9	12.4	13.7	14.8	16.1
Y50-54	12.1	11.8	13.1	13.2	15.8	14.5
Y55-59	14.6	10.4	12.0	11.4	15.1	10.4
Y60-64	10.6	7.1	5.6	7.8	8.4	5.5
EDU						
No formal education	0.0	0.0	0.3	4.2	0.1	1.0
100 Primary	1.7	13.0	1.0	5.1	1.3	5.8
200 Lower secondary	10.3	19.8	8.6	17.4	22.9	34.0
300 Upper secondary	38.8	27.1	43.0	30.8	46.3	43.6
400 Post-secondary non-tertiary	15.1	10.3	0.1	0.1	2.1	1.3
500 Short-cycle tertiary	0.6	0.2	16.5	9.6	0.1	0.6
600 Bachelor's or eq. level	19.9	15.7	14.1	12.1	7.3	5.2
700 Master's or Ph.D	13.5	13.8	16.3	20.8	19.9	8.6
FTPT						
1. Full-time job	71.1	71.6	83.6	79.6	83.2	77.5
2. Part-time job	28.9	28.4	16.4	20.4	16.8	22.5
TEMP						
1. Permanent job	81.8	78.3	74.7	69.8	66.8	68.0
2. Fixed-term job	10.5	14.0	13.2	17.3	12.1	17.7
3. SELF-EMPLOYED	7.7	7.7	12.0	12.9	21.1	14.4
NUMJOB						
1. Only one job	95.5	94.4	94.8	93.7	98.8	98.5
2. Two jobs	4.2	5.3	3.9	4.1	1.1	1.3
3. Three jobs	0.3	0.3	1.3	2.2	0.1	0.2
Usual weekly working hours						
mean (SD)	35.1 (7.1)	34.3 (8.0)	37.4(10.6)	35.5(11.5)	37.6(3.8)	36.8(3.9)
Tenure (years in current job)						
mean (SD)	11.9 (7.1)	7.1 (5.8)	10.5 (5.7)	8.1 (9.1)	13.1(4.1)	8.2 (2.8)
Years from highest Educ. title						
mean (SD)	20.4 (8.7)	20.1(8.9)	19.9 (13)	20.7(13.7)	24.5(4.8)	25.3(4.2)

Table 1 (continued)

Individual variables (column %)	%	%	%	%	%	%
OCCUPATION (ISCO08_1D)						
1. Managers	4.6	3.1	6.9	6.1	3.8	2.0
2. Professionals	24.5	18.1	24.4	21.1	17.0	5.2
3. Technicians & assoc. profess	21.4	14.0	19.7	11.5	19.6	7.2
4. Clerical support	14.4	8.5	8.3	6.1	14.1	4.8
5. Service and sales	13.4	15.1	14.0	18.5	16.2	23.7
6. Skilled agricult./forest	1.4	0.7	2.8	1.6	2.1	1.9
7. Craft and related trades	10.4	12.3	9.3	12.2	12.0	19.2
8. Plant/machine operat./ assembl	5.0	10.3	6.4	6.8	6.3	9.4
9. Elementary occup	4.5	17.8	7.4	16.0	7.8	26.4
SECTOR (NACE2_1D)						
A-Agriculture, forestry and fishing	1.3	0.5	2.6	1.3	3.1	5.5
B-Mining and quarrying	0.2	0.1	0.1	0.1	0.1	0.1
C-Manufacturing	18.0	20.3	11.8	8.3	18.3	20.4
D-Electricity, gas, steam and air cond	1.1	0.5	0.8	0.5	0.7	0.2
E-Water supply; sewerage, waste	0.7	0.5	0.9	0.8	1.1	0.7
F-Construction	6.1	7.6	6.1	10.0	5.9	10.0
G-Wholesale and retail trade; repair	12.9	12.8	12.8	11.3	14.1	10.0
H-Transportation and storage	4.1	7.3	5.3	5.7	4.9	6.1
I-Accommodation, food service act	2.4	7.4	3.5	7.7	5.8	10.3
J-Information and communication	4.0	4.4	3.6	4.1	3.7	1.4
K-Financial and insurance act	3.6	1.6	3.4	2.7	3.0	0.5
L-Real estate act	1.0	0.7	1.2	1.1	0.7	0.4
M-Professional, scientific and technical act	5.7	4.4	6.4	7.0	7.4	2.3
N-Administrative, support service act	3.7	8.1	3.8	8.1	3.9	5.8
O-Public admin. and defence;	8.7	2.6	9.0	4.5	5.8	0.9
P-Education	7.5	4.3	8.2	6.4	7.9	2.4
Q-Human health and social work act	14.3	12.4	14.8	12.5	8.5	5.7
R-Arts, entertainment and recreation	1.3	1.1	2.2	2.0	1.5	1.0
S-Other service act	3.0	3.0	2.6	2.7	2.8	3.7
T-Act. of households as employers; act. of households for own use	0.2	0.5	0.7	2.9	0.8	12.4

while migrant households are more frequently family-based. Migration motives vary: family reasons dominate (57.8% in France, 49.7% in Italy), labour migration is highest in Italy (38.6%), and Germany has the largest share of asylum-related migrants (18.8%).

Section A2 in the Appendix presents and comments descriptive statistics.

Table 1 compares employment characteristics between nationals and migrant/foreign workers, also focusing on key occupational indicators. The profiles emerging show distinct national labour market configurations and contrasting positions of migrants within them.

Germany, France, and Italy show clear differences in migrant labour market integration. Germany has the strongest outcomes, with high employment (79.6%) and low unemployment (2.6%) among nationals, and relatively small gaps for migrants (72.8% participation; 5.2% unemployment). France shows weaker integration, with migrants having lower participation (70% vs 74.4%) and higher unemployment (11.6% vs 6.8%). In Italy, migrants participate more than nationals (71.2% vs 65.9%) but face the highest unemployment (10.4% vs 7.3%). Migrants are generally less educated, especially in Italy (only 5% tertiary), while France (20.8%) and Germany (13.8%) have more highly educated migrant shares.

Job quality and occupations differ markedly. In Germany, migrants are concentrated in manual jobs (≈ 40%) but have relatively stable employment conditions. France shows more segmentation, with migrants more often in temporary (17.3%) and part-time jobs (20.4%), mainly in services. Italy is the most polarized, with about 70% of migrants in low-skilled jobs and high shares in domestic work (12.4%) and hospitality (10.3%). Overall, Germany shows better integration, France more contractual inequality, and Italy stronger concentration in low-skill, precarious work.

Among the unemployed (not shown), migrants face higher unemployment than nationals in all countries, but with different patterns: Germany shows lower rates (5.2% vs 2.6%) and short durations, while France (11.6% vs 6.8%) and Italy (10.4% vs 7.3%) experience wider gaps and longer spells, especially in Italy where nearly half remain unemployed over six months. Inactivity (not shown) also differs, being more linked to education and health in Germany, while in France and especially Italy it is more often associated with care and family responsibilities (up to 16% among migrants in Italy).

These findings provide descriptive insights into the segregation or overrepresentation of migrant and national workers in certain labour conditions.

4 Methodology

Traditional analyses of gaps between migrants and nationals (among others Dustmann and Frattini 2013; Fasani and Mazza 2020, 2023; Kerwin and Warren 2020; Breem and Liebig 2020; Broberg et al. 2024) typically regress employment status—or a single outcome such as working in a specific sector or occupation—on a migrant–national dummy. However, this framework does not capture how a broader set of labour market characteristics (e.g., occupation, contract type) shape migrant–national disparities.

We use an approach that, treating migrant–national worker status as a binary outcome, offers a more natural framework for jointly assessing disparities across multiple labour market dimensions.

Specifically, the advantages of this approach are threefold. First, it allows us to account for a variable that is conceptually measured with error, in the sense that unobserved or unavailable characteristics (captured in the residual term) may differentiate migrant and native workers. This is consistent with the main focus of the study, which is to assess whether, conditional on observed characteristics, there remain unexplained differences in the probability of being a native or a migrant worker. Second, it enables the simultaneous modelling of the factors that distinguish native and migrant workers while explicitly accounting for selection effects (i.e. self-selection into the working group). Third, it allows us to jointly assess multiple labour market characteristics along which migrants and natives may differ, rather than analysing each outcome separately. In this sense, segregation in a given labour market condition (e.g. occupations), essentially measured as differences in the probability of being a migrant versus a native worker, is estimated while accounting for other job characteristics (e.g. sectors, number of jobs, etc.) and their potential segregation patterns.

In this perspective, one method for dealing with selection bias is to use a bivariate binary selection model (Heckman 1979) which controls for non-random sample selection by using a two-equation system, where one equation, for subject i, describes the selection process (dummy y_{i1}) and the other describes the outcome of interest (dummy y_{i2}). More specifically, y_{i1}=1 if subject *i* is employed (0 otherwise), whereas for employed people, y_{i2}=1 if the subject is a migrant worker (0 if national).

Using the latent variable representation, Y_{i1}^* is the random variable representing attitude to work, whereas Y_{i2}^* is a random variable representing the attitude to be a migrant worker. The labour market participation mechanism is represented by the Bernoulli variable y_1, and derives its properties from those of Y_{i1}^*: y_{i1}=1 is observed when Y_{i1}^* is higher than a certain arbitrary threshold (Y_{i1}^*>0), and analogously for Y_{i2}^* and y_{i2}.

Since we can model y_{i2} if and only if y_{i1}=1, the subsample with y_{i1}=0 is not of interest for assessing ethnic gaps in labour market outcomes (outcome equation). However, it is retained to control for the selection mechanism (selection equation).

Both equations could be modelled using a bivariate Probit model (Heckman 1979), where both selectivity issues are jointly determined in a double selection framework, jointly modelling two binary dependent variables (y_1 and y_2) and manifestations of underlying correlated latent variables (Y_1^* and Y_2^*). Within the GJRM framework, we specify a bivariate Logit model, composed of three equations:

$$Y_{1i}^* = \boldsymbol{\pi}'\mathbf{z}_i + \gamma_i \tag{1}$$

$$Y_{2i}^* = \boldsymbol{\beta}'\mathbf{x}_i + \varepsilon_i \quad \text{when } y_{1i} = 1 \tag{2}$$

$$\mathbf{F}(\gamma_i, \varepsilon_i) = \boldsymbol{C}(\varepsilon_i, \gamma_i, \rho) \tag{3}$$

where π is a vector of q parameters related to z_i covariates in the first equation (x and z may contain common covariates), x_i are covariates of the second equation (outcome equation) with parameters β.

The error terms γ_i and ε_i are assumed to be independent of their regressors z_i and x_i, and follow a bivariate general cumulative distribution function (cdf) F, where ρ is the errors' correlations.

The specification of the GJRM joint model rests on the fundamental Sklar's (1973) theorem that affirms that every two-dimensional cdf for two random variables ε_i and γ_i can be expressed as $F(\varepsilon_i, \gamma_i) = C(F_1(\varepsilon_i), F_2(\gamma_i); \rho)$ in terms of two one-dimensional marginals $F_1(\varepsilon_i)$, $F_2(\gamma_i)$, linked by another bivariate cdf C, called copula, allowing an association parameter ρ linking the correlation of two margins (Trivedi and Zimmer 2007). Our hypothesized error structure in Eq. (3) is to choose two logistic marginals, F_1 and F_2, for each error distribution, a Gaussian bivariate copula C where ρ is the correlation of two error terms.

Essentially, our specification is a joint system of two logistic regressions, where error terms of two equations are correlated using a bivariate normal distribution.

The significance of the ρ estimate determines whether the equations should be estimated jointly (if ρ is statistically significantly different from 0) or, alternatively, separately. In the latter case, it is effectively assumed that the observed sample of workers is not affected by self-selection, and inferences about the migrant–native gap can be drawn by analysing only the sample of employed individuals. Note that only when the null hypothesis cannot be rejected is it justified to estimate migrant–native differences using the sample of employed individuals, thereby ignoring selection.

The advantage of a bivariate logit model lies in the interpretability of its parameters. Once exponentiated, the estimated coefficients can be interpreted as classical odds ratios): for the first equation, π describes the impact of covariates on the attitude (odds) to be employed; for the second, β describes the impact of covariates on the attitude (odds) to be a migrant worker.

The above model was estimated using the GJRM R package (Marra and Radice 2017).

As mentioned above, classical analyses have measured the migrant–national gap by investigating the effect of a migrant/national dummy on a particular outcome variable.

In contrast, since our specification treats migrant/national worker status as the dependent variable in the outcome equation, our approach enables a comprehensive assessment of differences between both workers across all specified demographic and labour market characteristics, provided they are statistically significant in the second equation.

To enhance interpretability, beyond traditional Odds ratios—which only allow comparisons between each level of a categorical covariate and a reference category—we also present expected ratios (ER*). This measure provides insight into how the ratio among (percentage distribution of) expected migrants and expected nationals vary across different levels of a given covariate.

Specifically, let $P^*(y_{2i}=1 \mid X_i)$ and $P^*(y_{2i}=0 \mid X_i)$ represent the estimated probabilities for subject i of being a migrant or a national worker, respectively, given a set of characteristics X_i.

By summing these probabilities across N_g subjects in category g, where N_g is the sum of nationals and migrants in this category, ($N_g = NAT_g + M_g$) of a covariate x ($x_i = g$), such as age group or sex, we obtain the expected number of migrants

($M_g^*=\Sigma_{i\in Mg}\ P^*(y_{2i}=1\ /X_i,\ x_i=g)$ and national workers ($Nat_g^*=\Sigma_{i\in NATg}\ P^*(y_{2i}=0\ /X_i,\ x_i=g)$ within that category.

Aggregating these expected workers across all *G* categories and dividing by total expected workers in each group, we obtain percentage distribution for migrants $\mathbf{M_{\%}}^*=(M_1^*\ /\Sigma_g M_g^*,\ldots,M_g^*\ /\Sigma_g M_g^*,\ldots,M_G^*\ /\Sigma_g M_g^*)$ and nationals $\mathbf{NAT_{\%}}^*=(NAT_1^*\ /\Sigma_g\ NAT_g^*,\ldots,\ NAT_g^*\ /\Sigma_g\ NAT_g^*,\ldots,\ NAT_G^*\ /\Sigma_g\ NAT_g^*)$. To assess the relative expected number of migrants in relation to national workers within that category, we define Expected ratios (ER*) for category *g* as the ratio:

$$ER^*(g)=\left[M_g^*/\sum\nolimits_g M_g^*\right]/\left[NAT_g^*/\sum\nolimits_g NAT_g^*\right]\quad g=1,\ldots,G \tag{4}$$

A large ER*(*g*)>1 suggests an unexpected concentration of migrants in category *g* (e.g., specific sectors, occupations, contract types) relative to nationals, each group with their own characteristics.

In the case of disadvantaged job categories (e.g., elementary occupations), could indicate a form of labour market segregation affecting migrant workers. Conversely, a large ER*(*g*)<1 implies the opposite.

When using observed counts of Migrants and Nationals (M_g, and NAT_g respectively) in the data, formula (4) provides observed ratios, indicated by $OR(g) = [M_g\ /\ \Sigma_g M_g]\ /\ [NAT_g\ /\Sigma_g\ NAT_g]$.

4.1 Observed, expected and systematic segregation

Instead of ratios, by computing the differences, and summing their absolute values over G categories (dividing by 2) this leads to Duncan dissimilarity index for each (labour) variables (*x*) of interest

$$D^*(x)=\left[0.5\times\sum\nolimits_g|M_g^*/\sum\nolimits_g M_g^*|-|NAT_g^*/\sum\nolimits_g NAT_g^*|\right]\times 100\quad g=1,\ldots,G \tag{5}$$

Large D*(*x*) values summarize possible form of concentration of migrants in the levels of the covariate *x,* using expected values based on covariates in the model. Specifically, the index is a measure of segregation or integration between two groups.

Applying formula (5) with the observed counts (M_g, and NAT_g) yields D(*x*), the *observed segregation* for migrants. A value of 0 indicates complete integration, while a value of 100 indicates complete segregation. Intuitively, D(*x*) is the minimum proportion of one group that would need to be reassigned to different categories to make the two distributions identical.

The difference between D(*x*) and D*(*x*) represents the unexpected component of segregation explained by the model covariates (*expected segregation*). Minimal differences between the two indicate not only that the estimation approach successfully reproduces the observed segregation, but also that group comparisons based on the model's expected counts are methodologically sound.

To this end, to compare our approach based on sample selection models with a naïve specification that models only the migrant worker equation, we also compute

the Duncan index using a simple logistic regression, D* LOGIT(x). This model estimates the migrant equation with the same covariates as the SSM in the first analysis but does not account for sample selection within the employed population.

Although comparing D(x) and D*(x) is a sound approach to assess migrant segregation, literature (Hellerstein and Neumark 2008; Åslund and Skans 2010; Glitz 2014) suggests that comparison should be based on a set of particular set of characteristics on which the researcher wants to condition.

To this end, given that available covariates may be partitioned in working variables W (sector, occupation, etc.) and individual characteristics S (sex, age, education, etc.) we fit a second GJRM model without working variables W in the second equation and, as above described using expected counts for both ethnic groups, we estimated the expected segregation index due only to individual characteristics, $D^{*S}(x)$. More specifically, following Glitz (2014), we measure *systematic segregation* by the scaled Duncan index:

$$D^{*S}(x) = [(D(x) - D^{*S}(x))/(100 - D^{*S}(x))] \times 100 \quad (6)$$

which provides a measure of effective segregation not explained by S, but only on effect of sorting of workers into W categories.

Specifically, in this analysis where the first equation remains the same as in the first analysis, the second equation involves following demographics (S): sex, age class, number of components in the household, if parents live in the same household, education, field of study and years passed since person obtained the highest level of education. Working variables for which we measure *systematic* segregations were sector, occupation, types of contract (full-time/part time, fixed term/permanent).

4.2 Systematic segregation by counterfactual distributions

Indices in Eq. (4–6), however, may be influenced by differences in the covariate profiles between the two groups. For example, even if large $D^{*S}(x)$ express segregation for migrants controlling for individual characteristics S, this type of control is model based (conditioning by covariates). Instead, other conditioning methods, such as matching (e.g. selection of migrants and nationals with the same characteristics) would provide better comparisons.

In this perspective, to enable a more reliable comparison, instead of matching we estimate *only* for migrant workers, the counterfactual distributions of native workers, which represent the expected number of national workers in each category of a given covariate (e.g., sector, occupation, contract type) under the hypothetical scenario that both migrants and nationals had the same characteristics of migrants (X_M). Then, observed segregation may be compared to the segregation calculated on shares difference between migrant and counterfactual nationals (at X_M values).

Using the previous arguments on ER*(g) and D(x) we calculated counterfactual ratios $ER^C(g)$ and counterfactual Duncan dissimilarity index $D^C(x)$, both using NAT_g^c, the counterfactual distribution of native with same characteristics (X_M) of migrants.

Using only covariates S in the second equation (yielding counterfactual expected ratios and the Duncan counterfactual index, denoted as $ER^{CS}(g)$ and $D^{CS}(x)$, respec-

tively), the difference between $D(x)$ and $D^{CS}(x)$—and specifically the corresponding (scaled) index of *systematic segregation*:

$$D^{CS}(x) = [(D(x) - D^{CS}(x))/(100 - D^{CS}(x))] \times 100 \quad (7)$$

providing a measure of segregation that controls only for S (and reinforced by the fact that comparisons were made at identical X_M values for both groups, where S is a subset of X_M), but attributable solely to the sorting effect in W. For further details, including links with classical counterfactual analysis, see Sect. A3 in the Appendix.

5 Results

5.1 Observed ethnic disparities by educational attainment

Table 2 highlights differences between migrants and natives in educational attainment, also summarizing their relative concentrations through the Duncan index. The most pronounced gaps emerge at the lower end of the education spectrum, where migrant shares substantially exceed those of natives—for instance, in Germany, the proportion of migrants with only primary education is more than seven times (7.2) that of nationals. By contrast, differences are far smaller among highly educated workers. These patterns are broadly consistent across countries.

With the exception of Italy, non-EU/EFTA migrants tend to resemble natives more closely than the overall migrant population, although these differences are generally minor, apart from a few cases—such as the higher share (5.5) of short-cycle tertiary education among non-EU/EFTA migrants in France. France also shows fewer disparities between migrants and natives within the highly educated segment of the population.

Table 2 Duncan index in education levels, and education ratios (Migrant workers shares in education level g/National workers shares in education level g), by country and type of migrants

		All migrants			Non-EU_EFTA migrants		
		DE	FR	IT	DE	FR	IT
Duncan index	EDUCATION	20.0	19.3	15.7	20.9	20.6	26.6
Education ratios[⊥]	Primary	7.2	4.7	4.4	6.7	4.6	4.4
	Lower sec	1.9	2.0	1.5	1.8	1.6	1.9
	Upper sec	0.7	0.7	1.0	0.7	0.9	0.8
	Post-sec. non-tertiary	0.7	1.4	0.6	0.8	0.7	1.1
	Short-cycle tertiary	0.4	0.6	2.2	0.5	5.5	0.6
	Bachelor's or eq. level	0.8	0.9	0.8	0.9	0.7	0.8
	Master's/PhD	1.0	1.3	0.4	1.1	0.4	1.4

[⊥]Education ratios highlight educational categories where migrant workers are overrepresented (ratio > 1), underrepresented (ratio < 1), or have a similar share to national workers (ratio = 1)

Table 3 Share of Non-EU/EFTA Migrant Workers with Tertiary Education by Occupation and Sector, and Tertiary Education ratios (Percentage of Non-EU/EFTA Migrant Workers with Tertiary Education in occup./sector *g* Divided by Percentage of National Workers with Tertiary Education in occup./sector *g*

	Share of Non-EU_EFTA migrant workers with tertiary educ			Tertiary education ratios* (Non-EU_EFTA vs Natives)		
ISCO08_1D (Occupations)	DE	FR	IT	DE	FR	IT
1. Managers	50.6	64.0	18.2	0.9	1.0	0.6
2. Professionals	80.0	94.9	77.5	1.1	1.1	1.0
3. Technicians & assoc. profess	32.9	62.1	30.9	1.0	1.1	0.8
4. Clerical support	26.7	44.1	28.8	1.6	1.0	1.2
5. Service and sales	14.3	21.3	9.7	1.3	1.3	1.3
6. Skilled agricult./forest	21.4	14.3	3.8	0.9	0.6	0.5
7. Craft and related trades	13.4	16.3	4.3	1.2	1.0	3.0
8. Plant/machine operators/assemblers	14.2	18.9	5.8	2.5	2.1	3.9
9. Elementary occup	10.6	9.5	5.2	2.7	1.2	2.4
SECTOR						
A-Agriculture, forestry and fishing	23.1	5.3	2.0	0.8	0.2	0.3
B-Mining and quarrying	50.0	50.0		1.8	1.0	0.0
C-Manufacturing	32.5	39.4	8.7	1.0	1.0	0.6
D-Electricity, gas, steam and air cond	66.7	84.6	0.0	1.4	1.6	0.0
E-Water supply; sewerage, waste manag	10.0	13.0	-	0.6	0.4	0.0
F-Construction	21.6	19.5	4.9	0.9	0.8	1.0
G-Wholesale and retail trade; repair	19.2	40.3	8.9	1.1	1.1	0.7
H-Transportation and storage	18.2	26.7	13.1	1.1	0.9	1.1
I-Accommodation and food service act	15.6	24.6	8.2	1.5	1.0	1.2
J-Information and communication	75.0	89.3	52.6	1.4	1.1	1.1
K-Financial and insurance act	61.7	89.7	75.0	1.7	1.1	1.6
L-Real estate act	35.0	43.8	25.0	1.2	0.7	1.2
M-Professional, scientific and technical act	66.9	86.3	69.0	1.2	1.1	1.2
N-Administrative and support service act	18.5	22.8	8.4	1.0	0.6	0.6
O-Public admin. and defence;	42.0	35.7	12.5	1.0	0.7	0.3
P-Education	64.8	62.3	51.7	0.9	0.9	0.8
Q-Human health and social work act	30.3	37.0	23.9	1.1	0.8	0.4
R-Arts, entertainment and recreation	51.9	54.3	31.6	1.4	1.0	0.9
S-Other service act	21.0	47.4	14.8	0.6	1.3	1.0
T-Act. of households as employers; act. of households for own use	8.3	19.2	8.5	1.0	1.7	2.6

*Tertiary education ratios highlight occupational/sectorial categories where the most educated migrant workers are overrepresented (ratio > 1), underrepresented (ratio < 1), or have a similar share to national workers (ratio = 1)

5.2 Observed ethnic disparities among tertiary-educated workers in occupations and sectors

For a deeper examination of ethnic disparities among the most educated individuals across key occupational dimensions, Table 3 summarizes differences in the distribution of tertiary-educated migrant and native workers across occupations and sectors.

Concerning occupations, a large proportion of migrant workers hold tertiary degrees in professional occupations—ranging from 77.5% in Italy to 94.9% in France—figures that are almost identical to those of national workers. By contrast, in elementary occupations, tertiary-educated migrants are nearly three times as prevalent as their native counterparts in Germany (observed ratio, Tertiary education ratio = 2.7), with 10.6% of migrants versus 3.9% of nationals holding a degree in this occupational group, with fewer disparities in France.

Strong concentration is observed among highly educated migrants employed in manual and operational occupations, particularly in the "Plant and machine operators/assemblers" category across all countries—most notably in Italy (Tertiary education ratio = 3.9)—and in "Craft and related trades" (Tertiary education ratio =3.0) in Italy. Sectoral patterns reveal similarly pronounced disparities: the highest segregation levels appear in "Activities of households as employers" in Italy (Tertiary education ratio =2.6) and France (Tertiary education ratio =1.7), as well as in "Financial and insurance activities" (OR =1.7) and "Accommodation and food service activities" (Tertiary education ratio =1.5) in Germany.

5.3 Observed ethnic disparities in occupations and sectors, by educational attainment

To summarize patterns of segregation—both overall and by educational attainment—across sectors and occupations, Table 4 (replicating Table 6 from Dustmann and Frattini 2013) reports overall migrant segregation as measured by Duncan indices. The table also distinguishes between different types of migrants.

The estimated dissimilarity indices confirm substantial cross-country and educational heterogeneity in the occupational and sectoral allocation of immigrants relative to natives. Germany and Italy display higher overall occupational segregation than France.

Table 4 Duncan dissimilarity index (D): overall, for aggregated education group, and weighted average index, by country and migrant type

		ISCO08_1D			SECTOR		
Duncan index	Country	DE	FR	IT	DE	FR	IT
Overall	EU_EFTA vs NAT	24.0	17.4	32.9	18.4	20.0	27.2
	Non-EU_EFTA vs NAT	18.9	16.6	39.8	16.8	15.3	29.7
EDU High (EDU ≥500)	EU_EFTA vs NAT	16.4	10.0	29.5	14.2	13.8	23.3
	Non-EU_EFTA vs NAT	11.7	10.6	37.9	18.6	13.2	32.3
EDU Medium (EDU 300–400)	EU_EFTA vs NAT	18.3	18.3	34.9	16.0	18.5	27.1
	Non-EU_EFTA vs NAT	14.9	14.3	37.4	12.4	16.4	26.9
EDU Low (EDU ≤200)	EU_EFTA vs NAT	26.3	22.7	13.0	24.7	36.0	21.6
	Non-EU_EFTA vs NAT	20.8	19.2	21.3	20.0	20.2	20.1
Weighted average across EDU	EU_EFTA vs NAT	19.0	15.1	27.7	16.8	18.5	24.7
	Non-EU_EFTA vs NAT	14.8	13.2	33.2	15.7	15.4	26.4

The table reports the Duncan dissimilarity index for the distribution of EU_EFTA (including Europe, EFTA) and non-EU_EFTA migrants versus natives, across 1-digit ISCO occupations (left), and across sectors (right**).** Rows 1–2 report the overall index. Rows 3–8 report the index by education group. Rows 9 and 10 report the average of the index by education group weighted by the share of total population in each education group

Occupational dissimilarity among EU immigrants reaches 24% in Germany and 33% in Italy, compared with a much lower 17% in France. For non-EU immigrants, the index is 19% in Germany, 17% in France and 40% in Italy, indicating that in the first two countries although segregation remains notable, these indices show comparatively integrated occupational structures. The pattern again places Italy as the country with the strongest segregation, whereas France exhibits the weakest. Sectoral indices are consistently smaller—around 18–20% in Germany and France, and 27–30% in Italy—confirming that disparities across industries are less marked than those across occupations, though still highest in Italy.

Across education groups, dissimilarity tends to rise as education declines, yet the gradient varies by country. Among the highly educated, indices are relatively low—around 10–16% in Germany and France, and 30–38% in Italy—showing near parity with natives in Germany and France but continuing gaps in Italy. Among the low educated, segregation intensifies: in Germany, EU migrants reach 26%, and in France roughly 23%, while in Italy occupational indices among EU migrants fall to 13% but remain elevated for non-EU workers (21%), suggesting that sectoral rather than occupational clustering dominates in Italy's low-skilled segment. Thus, Italy remains the most segmented labour market, followed by Germany, while France retains the most integrated structure—broadly consistent with the comparative hierarchy highlighted in Dustmann and Frattini's findings.

In this perspective, compared with their results for 2007–2009, our findings indicate that occupational segregation has eased in Germany and France but remains structurally high in Italy. The inversion in Germany—where non-EU immigrants now appear less segregated than EU ones—suggests gradual convergence through longer residence and improved credential recognition. France's small overall decline confirms its historical advantage linked to earlier and more diverse immigration inflows. Italy, instead, shows continuity with the earlier pattern highlighting that, despite limited reductions in sectoral indices, foreign workers—especially non-EU—remain concentrated in low-status occupations, reflecting a segmented and still maturing labour market.

To conclude this descriptive part, Fig. 7 in the Appendix also reports the Duncan indices across occupational levels by sex, age group, origin, and recency of migration. Overall, France exhibits lower levels of migrant segregation across almost all categories compared with Germany and, in particular, Italy. The most segregated groups are Non-EU women and Non-EU individuals aged over 50 in France and Germany (and those aged 30–50 in Italy), as well as African and Asian migrants in Italy, whereas France and Germany display lower and more homogeneous dissimilarity levels across ethnic groups. Segregation by recency of migration appears similar across countries. The least segregated groups are Non-EU migrants under 30 in France and Germany, and EU migrants under 30 in Italy.

5.4 Ethnic disparities with sample selection models

Before presenting the main results (based on the country-of-birth definition), we estimated several null GJRM models (taking only NUTS2 effect in both equations),

Table 5 Estimated correlations among residuals (ρ*) in the bivariate model (null) specification including only NUTS2 effects, for different definitions or subgroups of migrants

	NULL MODELS				FINAL MODEL
	MIGRANT (Country Birth definition)	MIGRANT (Citizenship definition)	MIGRANT (excluding EU-EFTA)	MIGRANT (including only <5Y residence)	MIGRANT (Country Birth definition)
DE	0.674***	0.671***	0.658***	0.637***	0.245***
FR	0.684***	0.680***	0.685***	− 0.549***	0.468***
IT	− 0.784***	− 0.791***	− 0.720***	− 0.588***	− 0.012***

Sign. codes: '***' 0.001, '**' 0.01, '*' 0.05, '.' 0.1

varying the definition of migrant status and considering different subgroups, to assess the sign and magnitude of sample selection effects characterising migrants.

The estimated correlation coefficient, or sample selection parameter (ρ*), from this null model represents the raw, unadjusted association between the probability of employment and the propensity to be a migrant worker, controlling only for geographical fixed effects. It therefore provides a useful benchmark for assessing the magnitude and direction of the unadjusted relationship, as well as the extent of selection pattern.

In the null models (Table 5), Germany (DE, ρ*= 0.674) and France (FR, ρ*= 0.684) exhibit strong positive correlations for the general migrant population, suggesting that unobserved traits (e.g., ambition, general ability) and omitted observable characteristics conducive to working are the same traits characterising migrant workers. Conversely, Italy (IT) exhibits strong negative selection (ρ*= −0.784), indicating that unobserved characteristics increasing the likelihood of employment tend to reduce the likelihood of being a migrant worker in the Italian sample. This suggests the presence of distinct selection mechanisms influencing the decision to enter employment, as opposed to the decision to work in the host country (and possibly even the decision to migrate).

Crucially, for recent migrants in France (residing <5Y), the null model shows a highly negative correlation (ρ*= −0.549). This profound contrast with the overall French migrant population suggests that the unobserved characteristics that increase an individual's likelihood of securing *employment* strongly decrease their likelihood of being a *recent migrant worker* relative to a native worker. The final model (after selection of significant covariates in both equations) shows that controls absorb most unobserved heterogeneity, especially in Italy (ρ*= −0.012, not significant), though positive selection remains pronounced in France (ρ*= 0.468).

5.4.1 Selection equation

Significant covariates in selection equations (Table A3 and Fig. 6 in the Appendix, Sect. A3) show how individuals aged 15–64 sort themselves into employment according to gender, age, education, and household composition, with distinct national patterns. Men are systematically more likely to be employed than women, with the gender gap most pronounced in Italy. Age effects are strong in all countries but much steeper in France and Italy, indicating a sharp decline in employment probabilities at older ages. Individuals in larger households tend to show higher employ-

ment probabilities, pointing to economies of scale or greater economic need within multi-member families. However, when parents live in the same household as the head, employment chances are significantly lower—especially in France and Italy—suggesting that extended family co-residence is associated with weaker individual labour market attachment.

Urban residence is positively linked to employment, while both the number of children and, in particular, the presence of children aged 0–2 years act as significant barriers to labour market participation. The length of time since obtaining the last educational title also reduces employment chances, signalling declining attachment to the labour market among those who completed education longer ago.

Education and field of study further differentiate countries. In Germany, individuals sort into employment relatively evenly across education levels and disciplines, reflecting a more balanced labour market structure. In France and Italy, by contrast, employment chances rise steeply with education, and specific fields such as education, health, and engineering are associated with stronger attachment to employment. France also exhibits significant regional (NUTS2) variation, pointing to marked territorial differences in labour market opportunities.

Regional covariates (exhibiting significant effects especially in Italy) further shape employment outcomes and highlight broader contextual influences. In Germany, higher regional wages strongly attract individuals into employment, confirming the role of wage incentives in shaping labour supply. Negative NEET effects in Germany and Italy indicate that regions with more dynamic youth employment or educational participation also attract more people into work, pointing to positive complementarities between youth activity and general labour market vitality. In Italy, higher early school-leaving rates correlate positively with employment probabilities, reflecting the persistence of low-skill job opportunities in local labour markets. These regional patterns show that local economic and educational environments significantly condition individual employment chances beyond personal characteristics.

Overall, Germany stands out as the most integrated system, supported by clear wage and labour incentives that foster employment participation. France displays stronger segmentation by education and age, reflecting the polarisation of employment opportunities along skill and generational lines, whereas Italy is characterised by deeper divides by gender, family composition, and regional disparities.

5.4.2 Outcome equation

Table A4 (Sect. A3 of the Appendix) reports significant covariates for the second outcome equation, which models the likelihood of being a foreign worker versus a native worker, conditional on employment. The results highlight both common and country-specific patterns in how individual and job-related factors shape migrants' positions within the labour market.

In Sect. A3 of the Appendix, we briefly discuss the selected instruments and link them to the relevant literature.

At the individual level, educational attainment and field of study are the main drivers differentiating migrant from native workers. Migrants are more likely to belong to lower educational categories, while the likelihood of being foreign declines with

higher education levels—an effect most pronounced in Italy, moderate in Germany, and weaker in France, where upper educational tiers show less differences. Field of study further accentuates these differences: migrants are concentrated in basic or vocational qualifications, manufacturing, and some service-related fields, while they are underrepresented in disciplines requiring higher credentials or public-sector credentials, such as education and law. This pattern is strongest in Italy, where migrants are clearly clustered in lower-skill domains, and less polarized in Germany and France.

The NUTS2 effects were highly significant across all three countries, confirming marked spatial heterogeneity in the distribution and integration of migrant workers. Regional covariates, by contrast, are generally not influential in Germany and France. In Italy, however, regions with a lower share of population with low educational attainment and a higher proportion of migrants in the population (Southern regions) show greater probabilities of migrant employment, suggesting that, at the local level, the coexistence of a large low-skilled labour segment and strong migrant presence tends to reinforce each other.

Overall, the results point to persistent but context-specific divides: Germany shows sharper occupational and contractual segmentation, France exhibits milder educational and occupational gaps, while Italy stands out for multidimensional disadvantage combining education, contract type, and sectoral concentration. For a clearer overview of covariates, Fig. 1 illustrates the mean probability of being a migrant worker across categories of the main educational and employment variables, whereas Fig. 2 shows the Expected ratios (Eq. 4), firstly considering all migrants (top figure) and then excluding EU and EFTA Migrants (bottom figure).

Figure 2 confirms previous results and reveals marked cross-country differences in the distribution of migrants and nationals across education and job characteristics. Migrants are consistently overrepresented in the lowest education levels—most notably in Germany, followed by Italy and France—and underrepresented at higher levels, with Italy showing the strongest educational segregation and France the weakest. Occupationally, migrants are concentrated in low-skill jobs (especially ISCO 9), with Italy and Germany displaying the highest ratios, while professionals and associate professionals (ISCO 2–3) are predominantly nationals, except in France where parity is greater.

Sectoral patterns confirm this divide: migrants are clustered in accommodation and food services (Germany), household employment—especially domestic work in Italy, where also Agriculture presents relevant migrant concentration— and underrepresented in public administration, education, and cultural sectors. In Italy, they are particularly scarce in professional and ICT-related occupations ($ER^* \approx 0.3$–0.4) and in industry ($ER^* = 0.4$), a pattern also observed in Germany ($ER^* = 0.5$). Nationals generally hold more permanent contracts, while migrants more often have fixed-term ones, with modest part-time differences except for slightly higher migrant part-time rates in Italy and France.

Excluding EU and EFTA migrants sharpens these disparities, especially in Italy, where non-EU migrants are even more concentrated in low-skill and domestic roles, while gaps narrow slightly in France. Overall, results highlight persistent multidimensional segmentation, particularly pronounced in Italy.

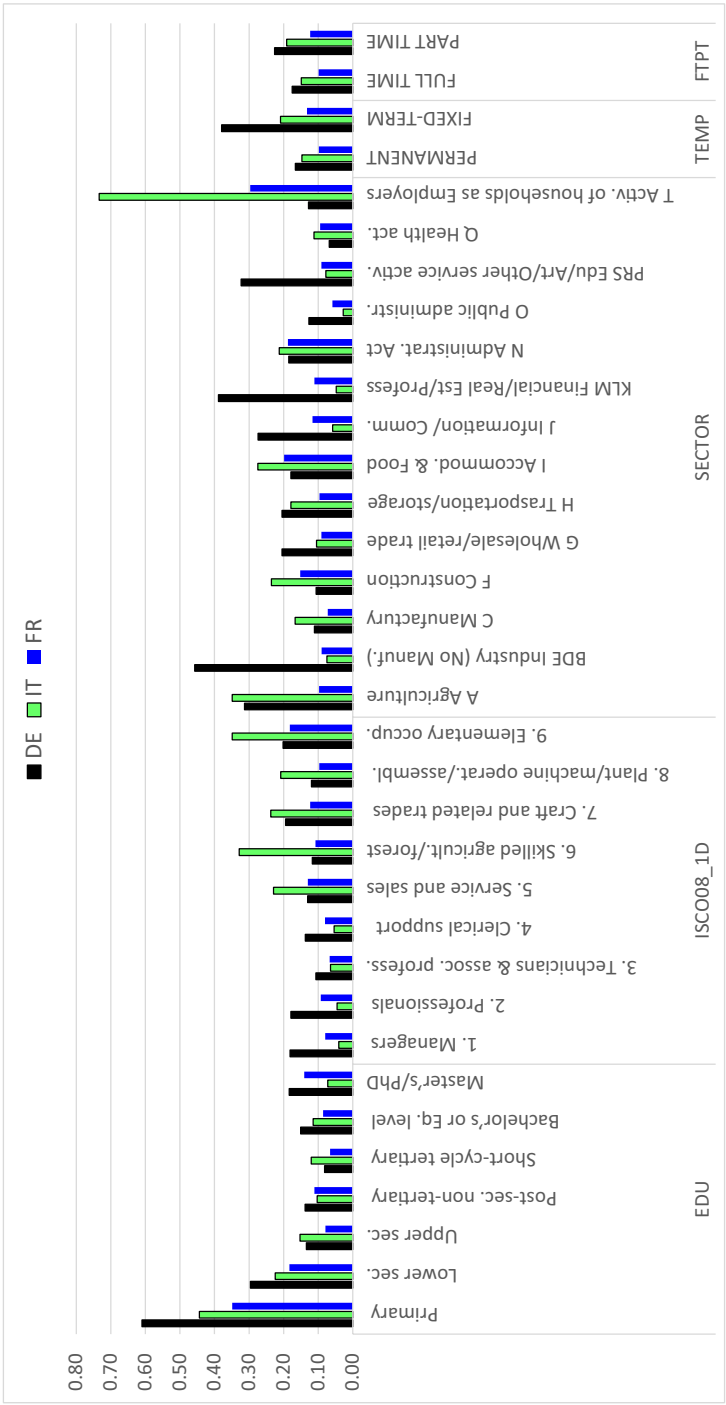


Fig. 1 Estimated mean probability to be migrant worker in categories of main variables (GJRM, second equation)

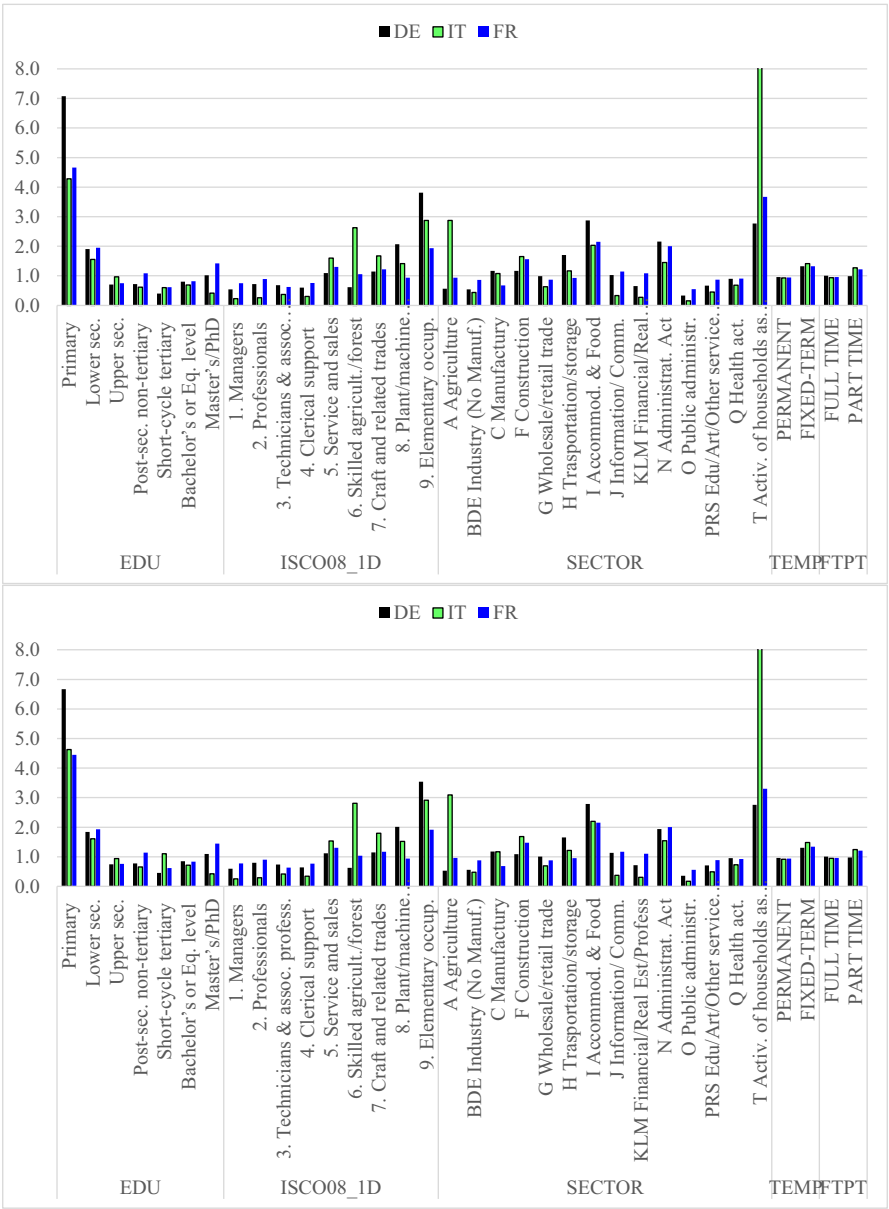


Fig. 2 Top: Expected ratios ER* (expected migrant worker shares / expected national worker shares) by education and working characteristics and country. Bottom: ER*, excluding EU and EFTA Migrants. For better readability, the values (top ER*=14.8; bottom ER*=15.2) for sector T in Italy were truncated at 8

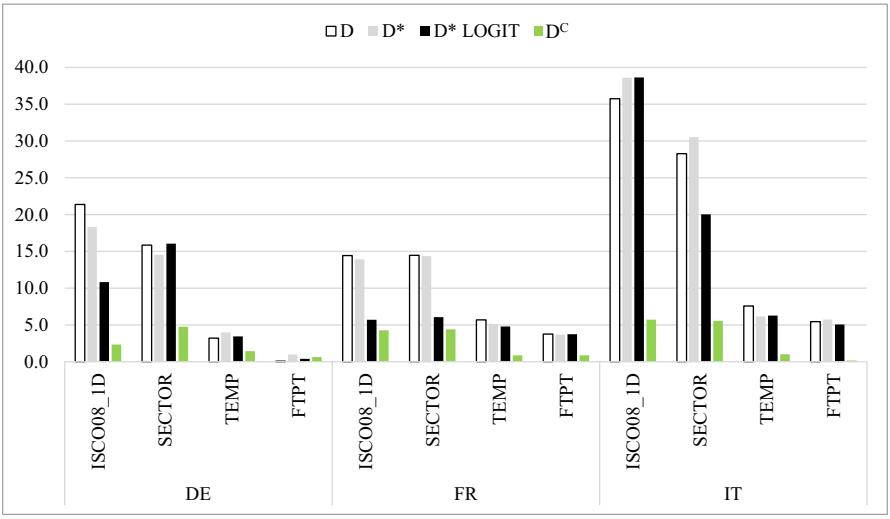


Fig. 3 Duncan dissimilarity index between Migrants and Natives: raw index (D), expected index under SSM (D*), expected index under logistic regression (D* LOGIT) and index with counterfactual natives (D^C). Expected indices and D^C were found using all individual characteristics (S) and working characteristics (W)

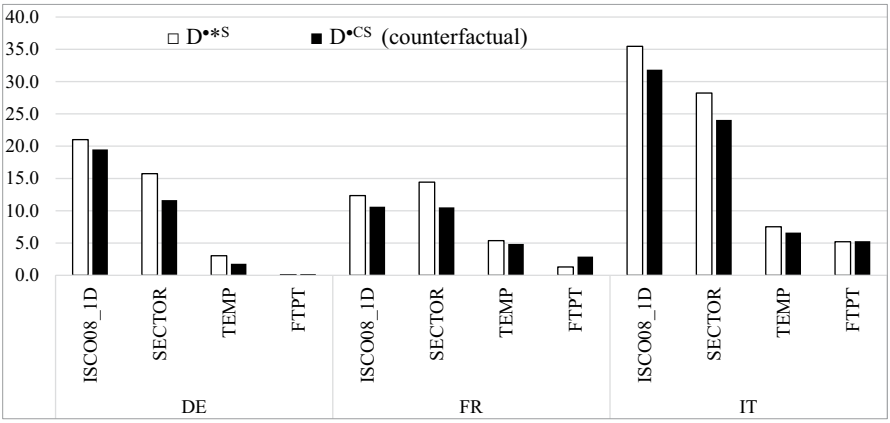


Fig. 4 Systematic segregation: Scaled Duncan index, controlling for individual characteristics S ($D^{\bullet *S}$) and scaled counterfactual Duncan index ($D^{\bullet CS}$) at same values of individual characteristics S of migrants

5.5 Segregation, expected segregation and systematic segregation

Observed and expected segregation (also found using counterfactual natives) were summarized in Fig. 3.

The *expected segregation* index D*, estimated using expected counts for migrants and natives by the full SSM model, closely matches the observed values D, unlike those derived from logistic regression models that ignore sample selection. As

expected, segregation levels decline substantially when counterfactual natives are used, since all covariates related to socio-demographic and working characteristics (*S* and *W*) are controlled for, and comparisons are made at the same covariate values for both groups.

However, *systematic segregation* (excluding the contribution of working variables from the expected migrants and expected nationals, while controlling for only S covariates) was measured by the scaled Duncan index ($D^{\bullet *S}$) proving the extent of migrant over representation relative to natives that depends solely on sorting across work-related categories (*W*).

Using a counterfactual approach, *systematic segregation* can also be measured by $D^{\bullet CS}$, comparing observed migrants and expected nationals (for the sample of migrants), at same levels of S covariates of migrants for both groups. Figure 4 demonstrates that systematic segregation persists, as evidenced by both measurement approaches.

Occupational systematic segregation ranges from about 35% in Italy to 13% in France, while sectoral segregation is slightly lower (28% in Italy and 14% in France). Apart from Italy, segregation related to contract type and working time is almost negligible.

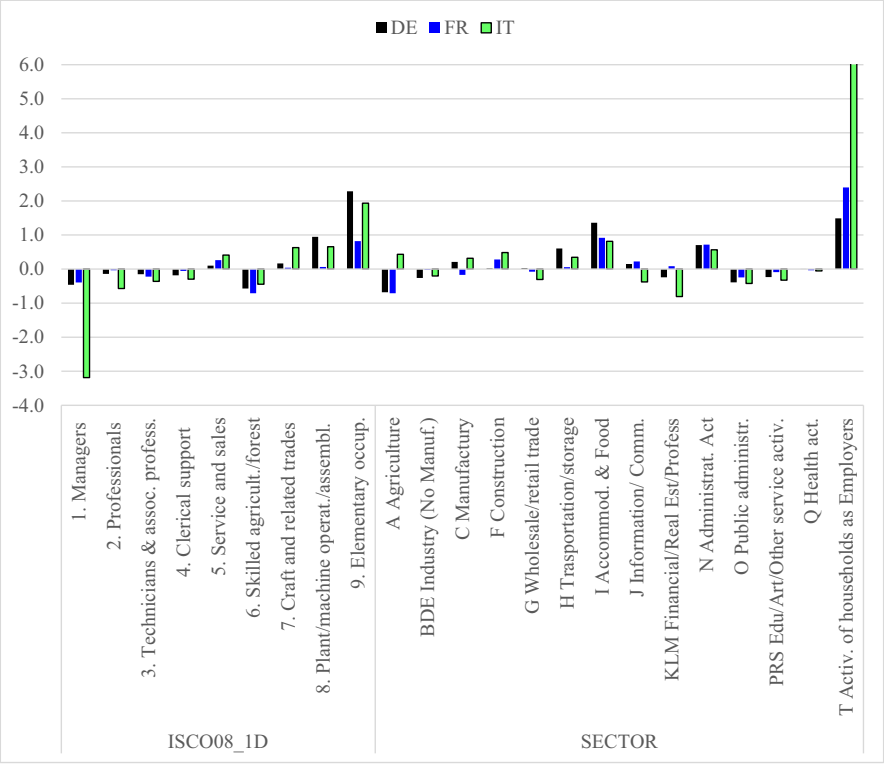


Fig. 5 Difference between the observed ratio OR(g) (migrant share / native share in category *g*) and the expected ratio $ER^{CS}(g)$ based on counterfactual nationals (at same levels of individual characteristics *S*). For better readability, the value ($OR - ER^{CS} = 13.1$) for sector T in Italy was truncated at 6

For a deeper investigation of these patterns, Fig. 5 shows in which specific categories systematic segregation is most pronounced, by comparing the observed ratios (raw shares of migrants and nationals) with the counterfactual ratios ER^{CS} (calculated at same levels of individual characteristics S of migrants).

Even when accounting for individual characteristics S and adopting a counterfactual approach, migrants remain concentrated in low-skilled and service-oriented positions.

Moreover, migrants are underrepresented in managerial and professional occupations, as well as in the financial, real estate, and professional sectors in Italy. Conversely, they are strongly overrepresented in elementary occupations—particularly in sector T in Italy (and, to a lesser extent, in France) and in sector I in Germany.

6 Conclusion

The findings indicate that Italian, French, and German national workers enjoy significant advantages in terms of job quality, access to high-skill occupations, and employment stability. In contrast, occupational and sectoral differences between highly educated migrants and nationals are minimal in certain areas—such as professional roles or sectors like ICT and professional, scientific, and technical activities.

The positive selection observed for Germany and France suggests that these destination countries offer higher returns to skills than the countries of origin. Such conditions may have incentivized high-skilled individuals to migrate, as they can benefit from better upward opportunities. This interpretation aligns with other evidence discussed above.

For example, Germany and France show better-educated migrant workers, stronger labour integration (France retains the most integrated structure), and smaller migrant–native gaps, while Italy combines high migrant participation with a strong concentration in low-skill, low-protection jobs. Migrants are disproportionately low educated, and the likelihood of being a foreign worker declines with education—most sharply in Italy, moderately in Germany, and only weakly in France.

Regional dynamics further reinforce these patterns. In Germany, higher wage levels are strongly associated with increased employment, reflecting a labour market that rewards skills and productivity. By contrast, Italy's compressed wage distribution makes employment more accessible to low-skilled workers. Moreover, in Italy, higher rates of early school leaving and lower average regional education levels are linked to higher employment—particularly among migrants—highlighting the persistence of low-skill labour markets alongside a large migrant workforce.

Regarding key disparities across ethnic groups, the share of tertiary-educated migrants employed in elementary occupations is more than twice that of nationals (2.7 in Germany, 2.4 in Italy), underscoring a clear pattern of migrant downgrading through overeducation, where workers hold qualifications exceeding the requirements of their positions. Only in France, however, do highly educated migrants and natives show a comparable concentration in elementary occupations, indicating that even well-qualified migrants may face partial skill underutilisation in certain seg-

ments of the labour market. This highlights the issue of migrant overeducation in Italy and Germany.

Migrant workers are heavily concentrated in low-skill occupations (e.g., elementary jobs) and particularly in Plant and machine operators or assemblers occupations and low-value-added sectors, such as household employment in Italy and accommodation and food in Germany.

This segregation, amplified by the higher proportion of tertiary-educated migrants in these concentrated professions, may indicate underlying discrimination. Moreover, the persistence of disparities between migrants and (counterfactual) nationals with same (identical) characteristics—particularly in Italy (where systematic occupational segregation is 35%)—suggests the presence of additional factors contributing to the unexplained advantages enjoyed by national workers over migrants.

While prior research has attempted to explain immigration-based segregation in the EU following theories based on gender segregation in occupation (e.g., Bettio and Verashchagina 2009; Carranza et al. 2019; Palencia-Esteban and del Río, 2024), it should be noted that migrants deal with additional (and often difficult to measure) sorting mechanisms that limit their occupational choices and outcomes.

Migrants encounter significant obstacles, such as legal restrictions (e.g., visa regulations, naturalisation policies) and language barriers (Bossavie et al. 2020; Broberg et al. 2024). Another major challenge is the difficulty in obtaining recognition for foreign qualifications, which contributes to migrant overeducation and leads to a substantial underuse of their tertiary education across the EU (Belmonte et al. 2020).

A key explanation for migrant overeducation is the imperfect transferability of credentials and skills acquired abroad, as supported by Chiswick and Miller (2008, 2009). One of the few studies examining this issue found that the strongest predictor of overeducation among migrant workers is having obtained their highest educational qualification outside the host country (Pecoraro and Wanner 2019).

The impact of foreign qualification recognition is evident in the segregation of migrants into elementary occupations and the unexpected overrepresentation of nationals in professional roles across all three countries. In Italy, this pattern is further reinforced by the marked concentration of tertiary-educated migrants in the "Activities of Households as Employers" sector, where their share is nearly three times that of natives.

Looking at observed and expected counterfactual ratios, in Germany and France tertiary-educated migrants do not face an "unexpected disadvantage" in professional occupations (ISCO 2) or in technician and associate professional roles (ISCO 3), where tertiary education is a key requirement.

Instead, since large under-representation emerges for tertiary-educated migrants on Skilled agriculture occupations (ISCO 6) and Agriculture sector in France and Italy, or industrial sectors (B, D, and E) in Italy and Germany, this may reflect the lack of recognition of foreign credentials in these specific segments of the labour market.

On the opposite, sectors such as manufacturing and ICT, concentrations appear broadly similar in France and Germany for both tertiary-educated workers, but also in general, suggesting that technical occupations or specific industries may help overcome barriers that persist in other segments of the labour market.

In jobs where practical skills and hands-on experience are more important than formal credentials or training certificates, the unexplained advantages for nationals and the factors driving migrant segregation appear to be less pronounced. This aligns with Pecoraro and Wanner (2019), who noted that while foreign qualifications and experience may not always be formally recognised, migrants can still secure employment based on the quality of their work.

Another motivation accords with Dustmann and Glitz (2015) and Lewis (2011), who proposed that analysing exogenous variation in the supply of immigrants to the labour market can provide causal evidence that firms adjust production methods to changes in the supply of skills.

From this perspective, a common finding is that migrant workers do not replace native workers, but rather accelerate the demand for everyday goods and services (Friedberg 2001; Dustmann et al. 2005). Two studies, one in the U.S. (Hong and McLaren 2015) and one in Germany on refugees (Auer and Götz 2024), found that 1 immigrant creates between 1.2 to 1.3 jobs in a local labour market, mainly due to a heightened demand for additional labour in service and social work, which are typically covered by female workers.

Other explanations for migrant segregation can be inferred from the findings.

Migrants may accept unattractive jobs due to better alternatives being unavailable, as well as the comparatively high wages they offer (Pioré, 1979). They may also be less concerned about job prestige or stigma, as their peer networks abroad could be kept uninformed of their employment (Fan and Stark 2011). Additionally, migrants often tolerate non-standard hours due to being single or having family abroad and are more likely to work in high-risk occupations (Fan and Qian 2017).

Research has shown that migrant recruitment frequently depends on referrals, offering employers a quicker and more efficient alternative to formal hiring processes (Drever and Hoffmeister 2008). This reliance on social networks can facilitate job matching while also reinforcing existing labour market patterns.

Recent evidence indicates that employers and the public are more receptive to hiring migrants in essential sectors, such as long-term care, food services, and cleaning. This may partly explain their overrepresentation in these critical industries (Allen et al. 2023).

Employers in such essential sectors often turn to migrant workers as a short-term solution to labour shortages, ensuring workforce continuity and addressing immediate demand gaps (Wright and McLaughlin 2024, Aldin et al. 2010).

Moreover, migrant workers may experience poorer working conditions and occupational health outcomes, wherein they face disparities in job security, exposure to hazardous work environments, and reduced access to occupational health services (Salvatore and Baglio 2018).

Migrants may also face greater pressure to accept or retain undesirable jobs, thereby influencing employer preferences. In particular, recently arrived migrants may take temporary or demanding jobs due to limited alternatives, financial obligations (Le Dé et al. 2016), or the need for career advancement (Jahn and Rosholm 2013). They may remain in such jobs longer if their residence status depends on employment (Hussein et al. 2011) or if they have weaker access to social safety nets (Verschueren 2016). Additionally, their greater geographic mobility in response to

economic opportunities further differentiates them from national workers (Basso and Peri 2020).

Finally, societal attitudes toward migration shape migrant employment outcomes. Discriminatory practices and public perceptions influence labour market participation and opportunities (Dorn and Zweimüller 2021; Bertrand and Duflo 2017). Conversely, more positive attitudes are crucial for successful labour market integration (Lee et al. 2022).

6.1 Policy implications and paper limitations

Our results highlight the importance of effective policy actions for enhancing the occupational integration of migrant workers. Key strategies include strengthening skills development, preventing occupational downgrading, and addressing origin-based disparities in the labour market World Migration Report 2024 (McAuliffe and Oucho, 2024).

Educational attainment and field of study emerge as the main factors differentiating migrant and native workers. Among the low educated, segregation intensifies across all three countries, suggesting that in Italy's low-skilled segment it is driven more by sectoral than occupational clustering—a pattern that is less evident in France and Germany. Differences in field of study further reinforce these patterns: migrants are disproportionately concentrated in basic or vocational tracks, as well as in manufacturing and selected service-related fields.

Enhancing migrants' skills through targeted initiatives, such as the European Commission's *Employers Together for Integration* programme, and addressing discrimination are crucial steps toward economic integration. Programmes focused on job-specific training and language acquisition can significantly improve employment prospects—a priority highlighted by various international organisations (Benton et al. 2014; OECD 2017).

However, given the evidence of migrant overqualification in certain sectors and occupations, alongside relatively low segregation among tertiary-educated individuals, policy should focus on addressing the undervaluation of foreign credentials and the barriers migrants face in obtaining formal recognition of their qualifications. Professions requiring licensing or credential recognition—such as professional and associate professional roles in Italy, skilled occupations in agriculture in France and Italy, and industrial sectors (B, D, and E) in Italy and Germany—pose particular challenges, as highlighted by our findings.

In the past, the EU has actively pursued policies aimed at streamlining the mutual recognition of higher education diplomas and professional qualifications across Member States. The Bologna Process, for example, standardizes and ensures the quality of higher education credentials by facilitating their comparability across European countries.

Hence, regulated professions, such as nurses, midwives, doctors, dentists, and pharmacists, benefit from an EU-wide system of automatic qualification recognition. This framework has enhanced labour mobility by allowing professionals to practice in different Member States without undergoing complex requalification procedures.

Hence, simplifying the recognition of non-EU foreign qualifications by reducing costs and bureaucratic delays, increasing transparency, and providing administrative support could significantly improve labour market outcomes.

To sum up, a more inclusive labour market, reinforced by positive public perceptions, can help reduce discrimination and improve career opportunities for migrants. Favour-

able societal attitudes may also drive policy changes that facilitate the hiring of foreign workers, including those with lower skill levels. Additionally, cultural and ethnic differences play a key role in shaping these dynamics, making public sentiment particularly influential in the employment outcomes of non-EU migrants.

Before concluding, we briefly outline the main limitations of this study.

First, our analysis does not cover an extended period but focuses only on the most recent year available in the LFS (2024). However, since the average job tenure for both migrants and nationals is no less than five years, we expect that our findings capture persistent segregation (controlling for selection) rather than short-term fluctuations in labour market dynamics.

Second, although the LFS is representative of a country's residents, surveys on migrants are inherently challenging. This is due to their high mobility and the fact that responses often depend on migrants' proficiency in the local language. This represents an additional source of selection bias that is not accounted for in the present paper.

Third, the analysis does not distinguish between first-generation immigrants (foreign-born) and second-generation immigrants (native-born with one or both parents born abroad). It also does not differentiate between immigrants who completed their highest level of education prior to migration and those who obtained it in the destination country. This limitation prevents us from assessing whether issues related to the recognition of educational qualifications acquired abroad have played a role in shaping occupational and sectoral distribution gaps.

Indeed, given that employment disparities are closely linked to income inequalities (Dustmann and Frattini 2013), the lack of comprehensive data on household income or family wealth prevents us from controlling for these factors in the employment selection process or from explaining gaps between the two ethnic groups at comparable levels of economic status.

Fourth, data limitations prevented us from incorporating or comparing a range of relevant contextual factors. The regional covariates do not capture key institutional dimensions—such as labour market regulations, sector-specific policies, welfare provisions for migrants, and integration policies—that may shape migrants' employment outcomes and levels of segregation. This constraint limits our ability to fully explain the observed segregation patterns, especially from a cross-country perspective.

However, as these public policies are typically set at the national level, it is challenging to include them in regional analyses involving a limited number of countries.

Despite these limitations, by adopting a solid methodological approach, we have provided a clear and comprehensive picture of migrant employment in the largest EU countries. The analysis illustrates both the strength and direction of self-selection driving the correlation between labour market sorting and the status of migrant workers, as well as the main individual and regional determinants of migrant–native gaps, with attention to occupational and sectoral segregation.

Appendix

Section A1: Regional determinants and literature

See Appendix Table 6.

From this perspective, the literature offers several relevant insights.

Regions with higher GDP and wages naturally attract migrants seeking better financial opportunities. Economic disparities between regions act as strong migration drivers, prompting individuals to relocate from more economically depressed areas to those offering better employment opportunities and higher wages (Borjas 2015; Grogger and Hanson 2011). Conversely, regions characterized by high poverty rates, a large concentration

Table 6 Regional determinants of employment and migration status (LFS regional statistics). Unless otherwise specified, data refer to 2023

Variables	Description
GDP	Gross domestic product (thousand Euro) at 2022 market prices
MEAN_H_INCOME	Average gross household income (2022)
WAGES	2022 wages and salaries in Euro per employee in full-timeequivalents
POVERTY RATE	% of people at risk of poverty relative to the total population
HIGH-TECH EMPLOY. RATE	% of people employed in knowledge-intensive sectors to all employed
SCIENCE_TECH EMPLOY.RATE	% of employed in science/technology or tertiary educ. to labour force
POPULATION OVER 65	% of population over 65 years
MIGRATION CHANGE RATE	Total net migration change 2023–2022
ACTIVITY RATE	% of economically active population of the total population
FEMALE EMPLOY RATE	Employment rates for Female (age group 15–64 years)
MALE EMPLOY RATE	Employment rates for Male (age group 15–64 years)
RATE ISCED ≤ 2	% of residents with lower than secondary education
RATE ISCED 3–4	% of residents with upper-secondary and post-secondary education
RATE ISCED 5+	% of residents with tertiary education
PART_TRAINING_EDU	% of residents attending formal education or training last year
NEET RATE	% of youth aged 18–29 not in employment, education, or training
EARLY_LEAVE	% of early leavers from education or training (age group 18 to 24 years]
AGR	Employment rates in Agriculture, Fish and Forestry (age-group 15–64 years)
WEB	% of Households that have internet access at home
IMM	% of Population with foreign and non-EU citizenship
EDU_TER_TOT	% of Population with Tertiary educational attainment (age group 25–64)
EDU_TER_F	% of women with tertiary achievement
FERT_RATE	Fertility rate

of workers in the agricultural sector, and elevated youth unemployment—particularly where a substantial share of young people is neither in education, employment, nor training (NEET) or consists of early school leavers—tend to drive migration abroad in search of better economic opportunities. This pattern aligns with research on economic push factors in migration (Cattaneo and Peri 2016; Dustmann and Okatenko 2014).

The presence of knowledge-intensive industries and employment opportunities in high-tech sectors plays a crucial role in attracting skilled migrants. Regions with a greater concentration of employment in science, technology, engineering, and mathematics (STEM) fields or better connected (internet access) regions appeal to migrants seeking career advancement (Docquier and Rapoport 2012; Kerr et al. 2016). Additionally, education levels shape migration patterns, as regions with higher shares of tertiary-educated individuals tend to attract skilled migrants, while those with lower education levels may experience emigration due to a lack of opportunities (Belot and Hatton 2012; Beine et al. 2008).

Demographic regional factors, such as ageing populations or fertility rates, influence labour demand. Regions with a higher proportion of elderly residents often face labour shortages, making them more attractive to younger migrant workers (Fokkema and de Haas 2015; Zaiceva 2014). Similarly, high regional labour force participation rates can signal economic dynamism, also increasing the demand for migrant labour (Peri 2016).

Additionally, female employment rates provide insights into gendered migration patterns, as gender-specific regional labour market opportunities can drive migration decisions (Donato et al. 2011; Leoncini et al. 2024). Moreover, the history of past migration trends and prevalence of migrant population provide direct measures of regional migration dynamics. The rate of migration change reflects regions' ability to attract and retain migrants, making it a key indicator of long-term migration patterns (Hanson and McIntosh 2016).

Section A2: Descriptive statistics

See Appendix Table 7.

The population distributions across Germany, France, and Italy are broadly balanced by gender, with men accounting for roughly half of the population (50.6% in Germany, 49.1% in France, 50.1% in Italy). Age structures, however, differ more markedly: Germany and Italy show a stronger concentration in older working-age groups, particularly 55–59 (12.7% and 12.9%) and 60–64 (11.8% and 11.2%), whereas France has a more even distribution and a relatively larger share in the younger brackets (10.1% aged 15–19).

Educational attainment varies substantially. Upper secondary education is predominant in all three countries but reaches its highest share in Italy (42.5%), while tertiary education (ISCED 5–8) is more prevalent in France (37.1%) and Germany (29.0%). Italy stands out for its concentration of individuals with lower secondary education (32.7%), contrasting with Germany (14.9%) and France (17.1%). Post-secondary non-tertiary and short-cycle tertiary levels are minor in Italy and France but represent 13% of the German population.

Citizenship and migration profiles reveal that nationals dominate everywhere, though the share of migrants (by citizenship) ranges from 17.7% in Germany to 10.1% in Italy and 7.8% in France. When considering birthplace, migrants account for 23.3% of the

Table 7 Descriptive statistics of individuals aged 15–64 by country: counts (Thousands) and Percentages

		DE		FR		IT	
		N	%	N	%	N	%
All People		53,677		40,399		37,143	
SEX	MALE	27,153	50.6	19,849	49.1	18,597	50.1
	FEMALE	26,524	49.4	20,550	50.9	18,545	49.9
AGE_GRP	Y15-19	3941	7.3	4075	10.1	2883	7.8
	Y20-24	4498	8.4	3777	9.3	2907	7.8
	Y25-29	4969	9.3	3601	8.9	2961	8
	Y30-34	5571	10.4	3848	9.5	3185	8.6
	Y35-39	5549	10.3	4035	10	3329	9
	Y40-44	5451	10.2	4178	10.3	3731	10
	Y45-49	4775	8.9	4108	10.2	4458	12
	Y50-54	5763	10.7	4300	10.6	4758	12.8
	Y55-59	6841	12.7	4239	10.5	4781	12.9
	Y60-64	6319	11.8	4240	10.5	4150	11.2
EDU (ISCED) 000	No formal education	32	0.1	567	1.4	208	0.6
100	ISCED 1 Primary	4375	8.1	983	2.4	1276	3.4
200	ISCED 2 Lower secondary	7994	14.9	6918	17.1	12,159	32.7
300	ISCED 3 Upper secondary	19,010	35.4	16,904	41.8	15,798	42.5
400	ISCED 4 Post-second. non-tertiary	6688	12.5	56	0.1	586	1.6
500	ISCED 5 Short-cycle tertiary	253	0.5	5157	12.8	49	0.1
600	ISCED 6 Bachelor's or eq. level	9023	16.8	4621	11.4	2235	6
700	ISCED 7–8 Master's or Ph.D	6303	11.7	5193	12.9	4832	13
FIELD of STUDY	NA/No specialisation (EDU<300)	12,401	23.1	8468	21.0	13,643	36.7
	Other	7058	13.1	8499	21.0	5973	16.1
	Business Admin	8140	15.2	8034	19.9	1794	4.8
	Engineering & Engineer. trade	6526	12.2	4536	11.2	2258	6.1
	Basic qualifications	3832	7.1	673	1.7	4580	12.3
	Health	3617	6.7	2562	6.3	1147	3.1
	Personal Services	2177	4.1	2204	5.5	1365	3.7
	Architecture & Construction	2187	4.1	1574	3.9	1127	3.0
	Business Admin & Law	1977	3.7			2792	7.5
	Education	2302	4.3	706	1.7	428	1.2
	Social, behavioural science	803	1.5	1329	3.3	834	2.2
	Arts	1067	2.0	892	2.2	911	2.5
	Manufacturing, Processing	1592	3.0	924	2.3	292	0.8
URBANIZATION	CITIES	21,396	39.9	18,270	45.2	12,891	34.7
	TOWN-SUBURB	22,297	41.5	8,100	20	18,395	49.5
	RURAL AREA	9,984	18.6	14,029	34.7	5,856	15.8

Table 7 (continued)

		DE		FR		IT	
		N	%	N	%	N	%
MIGRANT_CIT	NAT	44,155	82.3	37,248	92.2	33,379	89.9
(Citizenship definition)	MIG	9,522	17.7	3,151	7.8	3,763	10.1
MIGRANT_CB	NAT	41,182	76.7	35,125	86.9	31,908	85.9
(Country Birth definition)	MIG	12,495	23.3	5,275	13.1	5,235	14.1
MIGRANT_CB2	NAT	41,182	76.7	35,125	86.9	31,908	85.9
(Country Birth definition)	EU_EFTA	7,720	14.4	1,538	3.8	2,771	7.5
	Non-EU_EFTA	4,775	8.9	3,736	9.2	2,464	6.6
MIGRANT_RECENCY	NAT	41,182	76.7	35,125	86.9	31,908	85.9
	MIG OLD (≥5Y residence)	9,001	16.8	4,534	11.2	4,726	12.7
	MIG RECENT (<5Y residence)	3,494	6.5	740	1.8	508	1.4
HAS_CHILD	NO	34,541	64.3	22,664	56.1	22,937	61.8
	YES	19,136	35.7	17,735	43.9	14,205	38.2
JOB STATUS	Employed	41,391	77.1	27,618	68.4	22,835	61.5
	Unemployed	1,347	2.5	2,219	5.5	1,931	5.2
	Outside the labour force	10,939	20.4	10,563	26.1	12,376	33.3
Only migrants (Country Birth definition)							
MIGRAT. REASON	Employ. Found job before migr		14.0		6.9		7.2
	Employ. Hope find job		5.5		10.8		38.6
	Family reasons		42.1		57.8		49.7
	Education or training		7.0		12.3		2.1
	Retirement		0.0		0.1		0.0
	International protection or asylum		18.8		7.9		0.8
	Other		12.6		4.2		1.5

German population, 14.1% in Italy, and 13.1% in France. Germany also hosts the largest share of European (EU/EFTA) migrants (14.4%), while France shows a higher concentration of migrants from outside Europe (9.2%).

Urbanization patterns differ sharply: in France, 45.2% live in cities compared to 39.9% in Germany and 34.7% in Italy. Italy shows a clear suburban concentration (49.5%), while Germany has the most balanced distribution across settlement types. Family structure also differs, with a higher proportion of individuals with children in France (43.9%) compared with Germany (35.7%) and Italy (38.2%).

Among migrants, family-related motives dominate everywhere but are especially frequent in France (57.8%) and Italy (49.7%), while labour-related migration ("hope to find a job") is far more common in Italy (38.6%) than in the other two countries (10.8% in France, 5.5% in Germany). Germany stands out for the highest share of migrants who moved for asylum or protection (18.8%).

Labour market participation diverges substantially. Germany exhibits the highest employment rate (77.1%), followed by France (68.4%) and Italy (61.5%).

Correspondingly, Italy records the largest share of individuals outside the labour force (33.3%) and a relatively high unemployment rate (5.2%). Overall, the table reveals

marked cross-country contrasts in education, migration composition, and labour force participation, with Germany showing stronger labour attachment and higher migrant diversity, France a more urbanized and educated population, and Italy a suburban concentration and lower employment integration.

Household structures reveal clear national patterns and consistent migrant–native contrasts. In Germany, household composition is diversified but marked by a stronger family orientation among migrants: while single-person households are common in both groups (about 20%), migrant families with young children are considerably more frequent (26% vs 19% among nationals). This suggests a younger, family-based migrant population compared to an older, more individualised native one. In France, the pattern is similar but more pronounced: migrants display a distinctly family-centred profile, with nearly three in ten living as couples with children under 15 (28.5%, against 23.4% for nationals), and fewer in single or childless-couple arrangements. In Italy, household composition diverges sharply from the northern European model. Native households are dominated by multi-adult, childless structures (38.3%), consistent with extended family living or older age profiles, whereas migrant households show a stronger presence of nuclear families and multi-adult units with children (25.9% and 17%, respectively). Across all three countries, migrants emerge as more likely to live in younger, child-rearing families and less often alone or in childless partnerships—underscoring demographic renewal and stronger family dependency within migrant populations compared with the aging, more individualised native structures.

Section A3: Model results

See Appendix Tables 8 and 9 and Figs. 6 and 7.

Instruments

Regarding the instruments affecting the probability of employment without structurally determining migrant or native status, the key justification is that migrants and natives do not differ in these characteristics, which act as universal determinants of labour supply rather than markers of ethnic origin.

Household-related constraints—such as the number of children and the presence of infants increase the opportunity cost of working for any individual, regardless of background (Heckman 1979). Although family structures may differ, these factors represent common socio-economic constraints influencing labour supply without defining workers' ethnic characteristics (Palencia-Esteban 2022; Becker 1964).

Similarly, years since completing education capture labour market attachment and skill depreciation affecting the entire workforce (Parey et al. 2017), without necessarily distinguishing migrants from natives, whose differences instead stem from other issues linked to the recognition of foreign qualifications (Dustmann et al. 2016).

Regional variables—such as wage levels, NEET rates, and early school-leaving rates—reflect the economic and educational context in which both groups operate (Cattaneo and Peri 2016). These factors act as common "pull" forces shaping employment probabilities for all residents (Borjas 2015). Likewise, urbanization (cities vs. rural areas)

Table 8 GJRM selection equation (event = employed): Odds ratio and 95% confidence intervals

	DE			FR			IT		
	OR	95%CI		OR	95%CI		OR	95%CI	
N. of components 1 vs >4	1.10	0.92	1.33	0.86	0.71	1.06	1.07	0.93	1.24
N. of components 2 vs >4	1.08	0.93	1.25	0.86	0.73	1.00	1.03	0.91	1.18
N. of components 3 vs >4	1.14	1.00	1.30	1.09	0.95	1.25	1.06	0.95	1.19
N. of components 4 vs >4	1.25	1.12	1.41	1.20	1.07	1.34	1.14	1.03	1.26
SEX M vs F	1.68	1.59	1.78	1.50	1.41	1.59	3.31	3.13	3.50
SINGLE HOUSEHOLD HEAD	0.66	0.60	0.72	0.73	0.66	0.81			
BOTH PARENTS** vs NO ONE	0.72	0.62	0.83	0.39	0.33	0.46	0.45	0.40	0.50
ONLY FATHER vs NO ONE	0.70	0.55	0.90	0.45	0.35	0.59	0.49	0.39	0.61
ONLY MOTHER vs NO ONE	0.62	0.53	0.71	0.39	0.34	0.46	0.47	0.42	0.53
AGE Y15-19 vs Y60-64	0.78	0.58	1.04	1.53	1.32	1.79	0.16	0.13	0.20
AGE Y20-24 vs Y60-64	1.12	0.87	1.44	4.38	3.88	4.94	1.16	1.01	1.34
AGE Y25-29 vs Y60-64	1.49	1.19	1.87	7.46	6.62	8.41	2.85	2.51	3.24
AGE Y30-34 vs Y60-64	1.96	1.60	2.40	8.77	7.75	9.92	4.32	3.82	4.88
AGE Y35-39 vs Y60-64	2.20	1.84	2.65	9.57	8.45	10.85	4.52	4.00	5.11
AGE Y40-44 vs Y60-64	2.74	2.33	3.22	9.60	8.47	10.87	4.96	4.41	5.59
AGE Y45-49 vs Y60-64	3.33	2.87	3.85	10.25	9.05	11.60	4.94	4.43	5.51
AGE Y50-54 vs Y60-64	3.20	2.84	3.60	9.90	8.80	11.14	4.49	4.05	4.98
AGE Y55-59 vs Y60-64	2.59	2.36	2.85	6.66	5.98	7.41	3.34	3.02	3.68
YEAR_RESID From birth vs <5Y	3.73	3.41	4.08	2.69	2.32	3.12	2.39	1.96	2.93
YEAR_RESID ≥ 5Y vs <5Y	2.85	2.58	3.16	1.95	1.66	2.28	2.13	1.73	2.62
EDU 100 vs 700	0.58	0.13	2.58	0.06	0.02	0.22	0.07	0.01	0.39
EDU 200 vs 700	1.22	0.28	5.34	0.09	0.02	0.34	0.14	0.03	0.76
EDU 300 vs 700	0.72	0.64	0.81	0.38	0.34	0.42	0.41	0.37	0.46
EDU 400 vs 700	0.91	0.81	1.03	0.21	0.12	0.38	0.49	0.39	0.61
EDU 500 vs 700	1.08	0.72	1.62	0.72	0.63	0.81	0.38	0.19	0.76
EDU 600 vs 700	0.80	0.71	0.89	0.69	0.61	0.78	0.45	0.39	0.52
Field* Architec. Construc. vs SCI	1.23	0.96	1.56	1.20	0.98	1.47	1.31	1.01	1.70
Field Arts vs SCI	1.18	0.91	1.54	0.98	0.78	1.22	0.83	0.65	1.06
Field Basic qualifications vs SCI	0.53	0.42	0.67	0.80	0.63	1.01	0.73	0.59	0.91
Field Business Admin. vs SCI	1.46	1.18	1.82	1.21	1.03	1.41	1.06	0.84	1.33
Field Business Admin. Law vs SCI	1.65	1.28	2.11				1.20	0.95	1.50
Field Education vs SCI	1.59	1.26	2.02	1.68	1.26	2.23	1.39	1.02	1.90
Field Engineering vs SCI	1.46	1.17	1.83	1.27	1.07	1.51	1.23	0.98	1.56
Field Health vs SCI	1.52	1.21	1.91	1.51	1.26	1.82	2.27	1.74	2.94
Field Manufact. process. vs SCI	1.08	0.84	1.38	1.27	1.02	1.60	1.09	0.76	1.56
Field Other vs SCI	1.25	1.01	1.55	1.08	0.93	1.26	1.05	0.86	1.29
Field NA/No specialisation	0.45	0.10	2.03	2.41	0.63	9.30	1.37	0.24	7.74
Residence Cities vs Rural area	0.67	0.62	0.73	0.67	0.62	0.72			
Residence Town vs Rural area	0.84	0.78	0.90	0.75	0.69	0.81			
NUMBER of CHILDREN	0.82	0.78	0.86	0.85	0.81	0.89	0.92	0.88	0.97
NUMBER CHILDREN 0-2Years	0.61	0.55	0.66	0.80	0.73	0.89	0.84	0.75	0.94
YEARS SINCE EDU TITLE	0.97	0.97	0.98						
NEET_RATE	0.96	0.95	0.97				0.96	0.94	0.98
LOG_WAGES	3.34	1.89	5.90						
MIGRATION_CHANGE							0.98	0.96	1.00
FEMALE_EMPL_RATE							1.01	1.01	1.02

Table 8 (continued)

	DE		FR		IT		
	OR	95%CI	OR	95%CI	OR	95%CI	
EARLY_LEAVE					1.03	1.02	1.05
NUTS2 effect (Chi-sq.)			106.8	***			

Sign. codes: '***' 0.001, '**' 0.01, '*' 0.05, '.' 0.1. *field of study (reference = SCI, Social, behavioural sciences. **Parents live in the same household (No one the reference category). In Italics regional covariates

captures variation in job availability affecting all workers rather than specific ethnic groups (Grogger and Hanson 2011).

By contrast, variables common to both equations—gender, age, education level, field of study, and household size—capture both differences between migrant and native workers and their employment propensity.

In this framework, while household constraints and regional conditions influence the likelihood of labour market participation, they do not define migrant status. Instead, this distinction is shaped by differences in skills (general vs. country-specific), migration choices driven by international and regional skill-price differentials (Gould and Moav 2016), and unobserved traits such as ability, ambition, and motivation (Chiswick 1999; Bertoli et al. 2016).

Section A4. Deriving counterfactual distributions and dissimilarity indices

By selecting the estimated probabilities of being a national worker only for migrants (M), the general term $P^*(y_{2i}=0/X_i)$ can be rewritten as $P^*(y_{2i}=0/X_M, M)$, where X_M are covariate for migrants and this probability was only calculated for migrants.

Hence, $P^*(y_{2i}=0/X_M, M)$ provides a natural estimate of the counterfactual probability of being a national worker for the *i*-th migrant worker. From this perspective, the counterfactual analysis can be interpreted as conditional matching, where the *i*-th migrant is matched with themselves, and the probability of being a native worker is evaluated as if they possessed their own characteristics (x_{iM}).

By summing these probabilities over category *g* of a variable *x* (e.g., sector), $NAT_g^c = \Sigma_{i \in g} P^*(y_{2i}=0/X_M, M)$ gives the counterfactual expected number of national workers in category g, e.g. the hypothetical number of national workers in category *g*, assuming they had the same characteristics as migrants (X_M).

Aggregating this across all *G* categories of *x* yields the distribution of counterfactual expected national workers, $\mathbf{NAT^c} = (NAT_1^c,\ldots, NAT_g^c,\ldots, NAT_G^c)$, and its percentage distribution $NAT^c_{\%} = (NAT_1^c/\Sigma_g NAT_g^c,\ldots, M_g^c/\Sigma_g NAT_g^c,\ldots, NAT_G^c/\Sigma_g NAT_g^c)$, ensuring that all *G* terms sum to 1.

To this end, we construct fictitious labour market distributions for a variable *x* ($NAT^c_{\%}$), where the expected (counterfactual) distribution of national workers aligns exactly with the attributes of migrant workers (since $\mathbf{NAT^c}$ was calculated only for the migrant group).

By computing (counterfactual) ratios $ER^C(g) = (M_g/\Sigma_g M_g) / (NAT_g^c/\Sigma_g NAT_g^c)$ we can assess the extent to which the over- or underrepresentation (concentration) of migrants in certain categories, not explained by the selected (X_M chosen in the second equation) covariates. Since both groups have exactly the same values X_M of migrants, any remaining differences should highlight structural and unexpected sorting patterns.

Table 9 GJRM migrant working equation (event = migrant worker): estimated odds ratio and significance, by country. REF is the reference category for OR

		DE		FR		IT	
	(Variable)	OR	Sign	OR	Sign	OR	Sign
Sex	SEX M VS F	1.01		1.02		1.07	**
Age	AGE_GRPY15-19	REF					
	AGE_GRPY20-24	2.88	***	2.63	***	2.41	***
	AGE_GRPY25-29	5.67	***	4.95	***	4.10	***
	AGE_GRPY30-34	8.24	***	6.28	***	5.25	***
	AGE_GRPY35-39	11.15	***	7.94	***	6.93	***
	AGE_GRPY40-44	11.37	***	9.81	***	7.13	***
	AGE_GRPY45-49	12.71	***	9.96	***	5.60	***
	AGE_GRPY50-54	9.70	***	9.80	***	4.77	***
	AGE_GRPY55-59	7.70	***	9.19	***	3.95	***
	AGE_GRPY60-64	7.66	***	12.17	***	4.06	***
Number of components	N comp >4	REF					
	N comp1	0.49	***	0.40	***	0.60	***
	N comp2	0.51	***	0.37	***	0.39	***
	N comp3	0.63	***	0.52	***	0.42	***
	N comp4	0.69	***	0.50	***	0.47	***
Parents live in household	BOTH PARENTS	REF					
	ONLY FATHER	1.06		1.19		1.27	
	ONLY MOTHER	1.59	***	1.74	**	1.84	***
	NO PARENTS	3.44	***	3.09	***	3.45	***
Field of study	Architec. Construc	REF					
	Arts	1.22		0.42	***	0.68	
	Basic qualifications	5.42	***	2.41	***	4.34	***
	Business Admin	0.99		0.76	*	1.72	**
	Business Admin. Law	0.77	*			0.58	*
	Education	0.77	*	0.39	***	1.45	
	Engineering	1.31	**	0.87		1.53	*
	Health	1.45	***	0.77		1.93	**
	Manufact. process	1.16		0.57	**	3.68	***
	Other	1.39	***	0.98		1.15	
	Personal Services	1.22		0.74		1.06	
	Social, behav. sciences	1.36	*	0.96		1.14	
	NA/No specialisation	0.00		0.09		0.68	
Education	EDU 100	REF					
	EDU 200	0.58	***	0.46	***	0.37	***
	EDU 300	0.24	***	0.22	***	0.37	***
	EDU 400	0.30	***	0.26	*	0.28	***
	EDU 500	0.13	***	0.25	***	1.37	***
	EDU 600	0.41	***	0.36	***	0.46	***
	EDU 700	0.54	***	0.60	***	0.43	***

Table 9 (continued)

		DE		FR		IT	
	(Variable)	OR	Sign	OR	Sign	OR	Sign
ISCO08_1D	ISCO08_1D 10	REF					
	ISCO08_1D 20	1.30	**	1.15		1.14	
	ISCO08_1D 30	1.68	***	1.39	*	1.62	***
	ISCO08_1D 40	1.63	***	1.72	***	1.45	**
	ISCO08_1D 50	2.50	***	3.07	***	5.18	***
	ISCO08_1D 60	1.63	*	2.41	**	6.28	***
	ISCO08_1D 70	3.11	***	3.25	***	6.59	***
	ISCO08_1D 80	4.63	***	2.53	***	5.45	***
	ISCO08_1D 90	7.11	***	3.69	***	8.50	***
Sector	SECTOR A	REF					
	SECTOR BDE	0.85		0.77		0.23	***
	SECTOR C	1.77	**	0.73		0.43	***
	SECTOR F	1.74	*	1.35		0.56	***
	SECTOR G	1.59	*	0.83		0.28	***
	SECTOR H	2.22	***	0.77		0.52	***
	SECTOR I	4.07	***	1.73	*	0.74	***
	SECTOR J	2.17	***	0.80		0.33	***
	SECTOR KLM	1.40		0.94		0.28	***
	SECTOR N	2.17	***	1.35		0.50	***
	SECTOR O	0.67		0.50	**	0.16	***
	SECTOR PRS	1.16		0.83		0.41	***
	SECTOR Q	1.85	**	0.82		0.47	***
	SECTOR T	2.12	*	1.60		4.94	***
FTPT	PART-TIME VS FULL TIME	0.83	***	0.93		0.85	***
PERM	FIXED TERM VS PERM	1.53	***	1.57	***	1.10	***
HWACTUAL	HOUR ACTUAL WORK	1.00		1.00		1.01	***
YSTARTWK	YEAR START WORK	1.05	***	1.04	***	1.04	***
%POP WITH ISCED<2	RATE_ISCED <2					0.96	*
%MIGRANTS ON POP	IMM					1.07	***
NUTS2 Effect (Chi-sq.)		1544.1	***	835.2	***	260.7	***

Sign. codes: '***' 0.001, '**' 0.01, '*' 0.05, '.' 0.1

Specific NUTS2 effects were not shown. For labels see Table 1. In italics regional covariates

Hence, the Duncan dissimilarity index $D^C(x)$ based on the differences between observed migrant shares $M_g / \Sigma_g M_g$ and counterfactual native shares ($NAT_g^c / \Sigma_g NAT_g^c$)—calculated from probabilities estimated through sample selection models—summarizes the segregation of migrants across a labour market covariate x, evaluated at the common values of the selected individual characteristics X_M for both groups.

Using only covariates S in the second equation (yielding counterfactual expected ratios and the Duncan counterfactual index, denoted as $ER^{CS}(g)$ and $D^{CS}(x)$, respectively), the difference between $D(x)$ and $D^{CS}(x)$—and specifically the corresponding (scaled) index of systematic segregation, $D^{\bullet CS}(x) = [(D(x) - D^{CS}(x)) / (100 - D^{CS}(x))] \times 100$—provides a measure of segregation unexplained by S (and reinforced by the fact that comparisons were made at identical X_M values for both groups, where S is a subset of X_M), but attributable solely to the sorting effect in W.

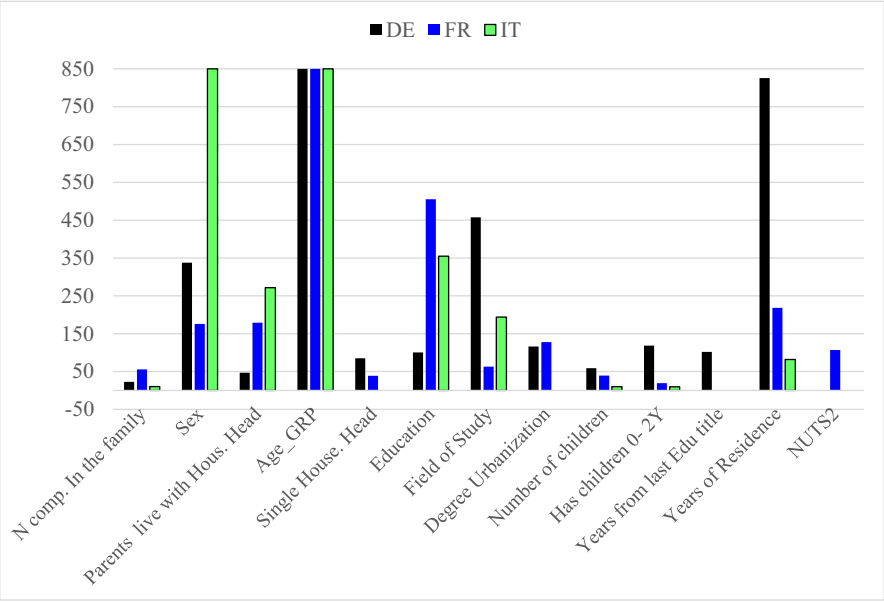


Fig. 6 GJRM selection equation (event = employed): Effects of individual covariates (Wald Chi-square sum of squares) and NUTS2 effects, by country. Age group effects (Wald Chi-square = 1096.1 in Germany, 3108.6 in France, and 2571.2 in Italy) and the sex effect for Italy (Wald Chi-square = 1735.7) were truncated at 850 for better readability of other effects

The proposed counterfactual approach is deeply rooted in the potential outcome framework of the Rubin Causal Model (RCM) (Rubin 2005). When treatment is not randomly assigned in non-experimental settings, $Y_i(1)$ and $Y_i(0)$ represent the potential outcomes for treated individuals (T=1, in this case, migrants) and untreated individuals (T=0, nationals), respectively. However, we can only observe either individual outcomes or their group averages (E), such as $E[Y_i(1) \mid T=1]$ for the treated and $E[Y_i(0) \mid T=0]$ for the untreated.

The difference in raw means between these two groups (measured by the raw dissimilarity index D in our setting) provides a biased estimate of the average treatment effect (ATE)—in this context, the overrepresentation of migrants compared to nationals across categories of a given (labour market) variable. This difference can be disaggregated into two components: the average treatment effect (ATE) for the treated/migrants ($ATE(T)=E[Y_i(1) - Y_i(0) \mid T=1]$) and a selection bias term,

$E[Y_i(0) \mid T=1] - E[Y_i(0) \mid T=0]$, which captures the pre-existing differences between the groups that are unrelated to treatment.

In our empirical framework, the counterfactual probability $P^*(y_{2i}=0/X_M, M)$ serves as an estimate of $Y_i(0) \mid T=1$ (the potential outcome if i-th migrant worker would be a national worker), while $\mathbf{NAT^c}$ provides an estimate of $E[Y_i(0) \mid T=1]$, whereas $D^C(x)$, the dissimilarity index among migrants and expected nationals, provides an estimate of ATE(T).

Hence, $D^C(x)$ allows for a comparison between the actual distribution of migrant workers and the expected distribution of national workers, assuming the latter had the same

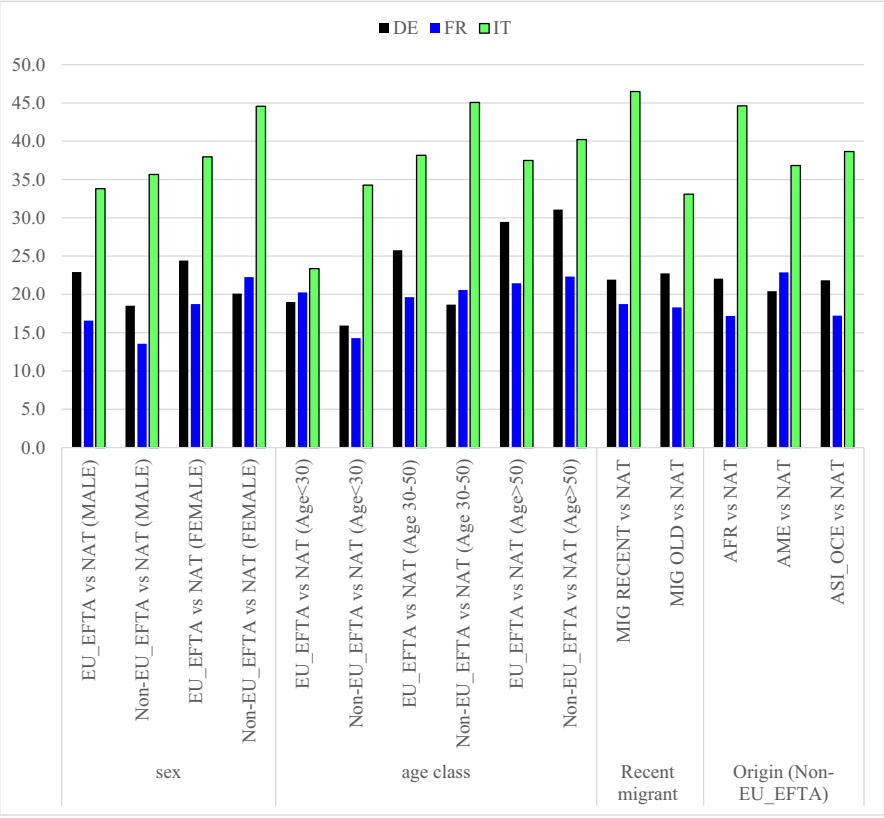


Fig. 7 Observed Duncan index, stratified by sex, age class, origin* and recency of migration**. * Origin for Non-EU_EFTA: AFRicans, AMEricans (North, South, Central), ASI_OCE (Asia, Oceania).** MIG RECENT are recent migrants (<5 years of residence in the host country), MIG OLD (≥5 years of residence in the host country)

characteristics as the former and specifically, the (chosen) covariates values observed for migrants (X_M).

Applying this logic, the difference among $D(x)$ and $D^C(x)$ expresses the contribution of selection bias (or unexplained sorting in categories of the x variable) to observed differences among migrants and nationals, thus systematic segregation.

Author contribution P.G.L wrote and reviewed the manuscript.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability The datasets analysed during the current study is available upon reasonable request.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Ethical approval The article does not contain any studies with human or animal participants and therefore, informed consent is not required.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Adsera A, Chiswick BR (2007) Are there gender and country of origin differences in immigrant labor market outcomes across European destinations? J Popul Econ 20:495–526

Aldin V, James D, Wadsworth J (2010) The changing shares of migrant labour in different sectors and occupations in the UK economy: an overview. In: Ruhs M, Anderson B (eds) Who needs migrant workers? Labour shortages, immigration, and public policy. Oxford University Press, Oxford, pp 57–85

Allen R, Pacas JD, Martens Z (2023) Immigrant legal status among essential frontline workers in the United States during the COVID-19 pandemic era. Int Migr Rev 57(2):521–556. https://doi.org/10.1177/01979183221127277

Angrist JD, Kugler AD (2003) Protective or counter-productive? Labour market institutions and the effect of immigration on EU natives. Econ J 113:F302–F331. https://doi.org/10.1111/1468-0297.00136

Åslund O, Skans ON (2010) Will I see you at work? Ethnic workplace segregation in Sweden, 1985–2002. Ind Labor Relat Rev 63(3):471–493

Auer D, Götz L (2024) Refugee migration, labor demand and local employment. Socio-Econ Rev 22(2):573–601. https://doi.org/10.1093/ser/mwad040

Ballarino G, Panichella N (2018) The occupational integration of migrant women in Western European labour markets. Acta Sociol 61:126–142

Bansak C, Simpson N, Zavodny M (2015) The economics of immigration, 1st ed. Routledge. https://doi.org/10.4324/9781315797250

Basso G, Peri G (2020) Internal mobility: the greater responsiveness of foreign-born to economic conditions. J Econ Perspect 34(3):77–98

Becker GS (1964) Human capital. National Bureau of Economic Research, New York

Beine M, Docquier F, Rapoport H (2008) Brain drain and human capital formation in developing countries: winners and losers. Econ J 118(528):631–652

Belmonte M, Grubanov-Boskovic S, Mazza J (2020) Foreign degrees, region of birth and under-utilisation of tertiary education in the EU (EUR 30041). Publications Office of the European Union, Luxembourg

Belot MV, Hatton TJ (2012) Immigrant selection in the OECD. Scand J Econ 114(4):1105–1128

Benton M et al (2014) Aiming higher: policies to get immigrants into middle-skilled work in Europe. Migration Policy Institute and International Labour Organization, Washington, DC

Bertoli S, Dequiedt V, Zenou Y (2016) Can selective immigration policies reduce migrants' quality? J Dev Econ 119:100–109

Bertrand M, Duflo E (2017) Field experiments on discrimination. Handbook of economic field experiments, vol Vol. 1. North-Holland, pp 309–393

Bettio F, Verashchagina A (2009) Gender segregation in the labour market: root causes, implications and policy responses in the EU. Publications Office of the European Union, Luxembourg

Biletta I et al (2019) How your birthplace affects your workplace. European Foundation for the Improvement of Living and Working Conditions, Luxembourg

Borjas GJ (1987) Self-selection and the earnings of immigrants. Am Econ Rev 77(4):531–553

Borjas GJ (2014) Immigration economics. Harvard University Press, Cambridge, MA

Borjas GJ (2015) Immigration and globalization: a review essay. J Econ Lit 53(4):961–974
Borjas GJ, Kauppinen I, Poutvaara P (2019) Self-selection of emigrants: theory and evidence on stochastic dominance in observable and unobservable characteristics. Econ J 129(617):143–171
Bossavie L et al (2020) Do immigrants push natives towards safer jobs? Exposure to COVID-19 in the European Union (Policy Research Working Paper No. 9500). Washington, DC: World Bank.
Breem Y, Liebig T (2020) How does migration shape industry structure? International migration outlook 2020. OECD Publishing, Paris. https://doi.org/10.1787/26a5b23b-en
Broberg N, Gonnot J, Poeschel F, Ruhs M (2024) Essential work, migrant labour: what explains migrant employment in European key sectors? OECD, ESPOL, Eurac, MPC, EUI
Brücker H, Glitz A, Lerche A, Romiti A (2016) Occupational recognition and immigrant labor market outcomes. www.sole-jole.org/16456.pdf. Accessed 14 June 2018
Carranza E, Das S, Kotikula A (2019) Gender-based employment segregation: understanding causes and policy interventions. World Bank, Washington, DC
Cattaneo C, Peri G (2016) What happens to the careers of European workers when immigrants "take their jobs"? J Hum Resour 51(3):655–693
Chiswick BR (1978) The effect of Americanization on the earnings of foreign-born men. J Polit Econ 86(5):897–921
Chiswick BR (1986) Is the new immigration less skilled than the old? J Labor Econ 4(2):168–192
Chiswick BR (1999) Are immigrants favorably self-selected? Am Econ Rev 89(2):181–185
Chiswick BR, Miller PW (2008) Why is the payoff to schooling smaller for immigrants? Labour Econ 15(6):1317–1340
Chiswick BR, Miller PW (2009) The international transferability of immigrants' human capital. Econ Educ Rev 28(2):162–169
d'Aiglepierre R, David A, Levionnois C, Spielvogel G, Tuccio M, Vickstrom E (2020) A global profile of emigrants to OECD countries: younger and more skilled migrants from more diverse countries. OECD social, employment and migration working paper No. 239. OECD Publishing, Paris. https://doi.org/10.1787/0cb305d3-en
De la Rica S, Glitz A, Ortega F (2015) Immigration in Europe: trends, policies, and empirical evidence. In: Chiswick BR, Miller PW (eds) Handbook of the economics of international migration. North-Holland, Amsterdam, pp 1303–1362
Docquier F, Rapoport H (2012) Globalization, brain drain, and development. J Econ Lit 50(3):681–730
Donato KM, Gabaccia D, Holdaway J, Manalansan M, Pessar PR (2011) Gender and international migration: from the slavery era to the global age. Russell Sage Foundation
Dorn D, Zweimüller J (2021) Migration and labor market integration in Europe. J Econ Perspect 35(2):49–76
Drever AI, Hoffmeister O (2008) Immigrants and social networks in a job-scarce environment: the case of Germany. Int Migr Rev 42(2):425–448. https://doi.org/10.1111/j.1747-7379.2008.00130.x
Dustmann C, Frattini T (2013) Immigration: the European experience. In: Card D, Raphael S (eds) Immigration, poverty, and socioeconomic inequality. Russel Sage Foundation, New York, pp 423–456
Dustmann C, Glitz A (2011) Migration and education. In: Hanushek E, Machin S, Woessmann L (eds) Handbook of the economics of education. Elsevier, Amsterdam, pp 327–439
Dustmann C, Glitz A (2015) How do industries and firms respond to changes in local labor supply? J Labor Econ 33(3):711–750
Dustmann C, Okatenko A (2014) Out-migration, wealth constraints, and the quality of local amenities. J Dev Econ 110:52–63
Dustmann C, Fabbri F, Preston I (2005) The impact of immigration on the British labour market. Econ J 115:F324–F341
Dustmann C, Frattini T, Preston IP (2013) The effect of immigration along the distribution of wages. Rev Econ Stud 80(1):145–173
Dustmann C, Schönberg U, Stuhler J (2016) The impact of immigration: why do studies reach such different results? J Econ Perspect 30(4):31–56
Eurostat (2023c) Foreign-born people and their descendants - Educational attainment level and skills in host country language. Eurostat statistics explained.
Eurostat (2022) Main characteristics of foreign-born people on the labour market. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Main_characteristics_of_foreign-born_people_on_the_labour_market
Eurostat (2024). Regional statistics. https://ec.europa.eu/eurostat/databrowser/explore/all/general?lang=en&subtheme=reg.reg_ilc&display=list&sort=category

Eurostat. (2023a). Migrant integration statistics: labour market indicators. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Migrant_integration_statistics_%E2%80%93_labour_market_indicators
Eurostat (2023b) Employment: annual statistics.https://ec.europa.eu/eurostat/statistics-explained/index.php/Employment_-_annual_statistics
Eurostat (2025) EU labour force survey database user guide (Version April 2025). Luxembourg: Publications Office of the European Union.
Fan W, Qian Y (2017) Native-immigrant occupational segregation and worker health in the United States, 2004–2014. Soc Sci Med 183:130–141
Fan CS, Stark O (2011) A theory of migration as a response to occupational stigma. Int Econ Rev 52:549–571
Fasani F, Mazza J (2020) Immigrant key workers: their contribution to Europe's COVID-19 response. IZA Policy Paper No. 155.
Fasani F, Mazza J (2023) Being on the frontline? Immigrant workers in Europe and the COVID-19 pandemic. ILR Rev 76(5):890–918. https://doi.org/10.1177/00197939231173676
Fernández-Reino M, and Sumption A (2024) Migrants in the UK Labour Market: an overview. Migration Observatory, University of Oxford. https://migrationobservatory.ox.ac.uk/resources/briefings/migrants-in-the-uk-labour-market-an-overview/.
Fokkema T, de Haas H (2015) Pre- and post-migration determinants of socio-cultural integration of African immigrants in Italy and Spain. Int Migr 53(6):3–26
Friedberg RM (2000) You can't take it with you? Immigrant assimilation and the portability of human capital. J Labor Econ 18(2):221–251
Friedberg RM (2001) The impact of mass migration on the Israeli labor market. Q J Econ 116:1373–1408
Glitz A (2014) Ethnic segregation in Germany. Labour Econ 29:28–40
Gould ED, Moav O (2016) Does high inequality attract high skilled immigrants? Econ J 126(593):1055–1109
Grogger J, Hanson GH (2011) Income maximization and the selection and sorting of international migrants. J Dev Econ 95(1):42–57
Grubanov-Boskovic S, Tintori G, Biagi F (2020) Gaps in the EU labour market participation rates: an intersectional assessment of the role of gender and migrant status. Publications Office of the European Union, Luxembourg
Hanson GH, McIntosh C (2016) Is the Mediterranean the new Rio Grande? US and EU immigration pressures in the long run. J Econ Perspect 30(4):57–82
Heckman JJ (1979) Sample selection bias as a specification error. Econometrica 47(1):153–161. https://doi.org/10.2307/1912352
Hellerstein J, Neumark D (2008) Workplace segregation in the United States: race, ethnicity, and skill. Rev Econ Stat 90(3):459–477
Hong G, and McLaren J (2015) Are immigrants a shot in the arm for the local economy? NBER Working Paper, no. 21123.
Hussein S, Stevens M, Manthorpe J (2011) What drives the recruitment of migrant workers to work in social care in England? Soc Policy Soc 10(3):285–298
Jahn E, Rosholm M (2013) Is temporary agency employment a stepping stone for immigrants? Econ Lett 118(1):225–228
Kaufman RL (2010) Race, gender, and the labor market: inequalities at work. Lynne Rienner Publishers, Colorado
Kerr SP, Kerr WR, Özden Ç, Parsons C (2016) Global talent flows. J Econ Perspect 30(4):83–106
Kerwin D, Warren R (2020) US foreign-born workers in the global pandemic: essential and marginalized. J Migr Human Secur 8(3):282–300. https://doi.org/10.1177/2331502420952752
Kugler AD, Pica G (2006) The effects of employment protection and product market regulations on the Italian labour market. In: Messina J, Michelacci C, Turunen J, Zoega G (eds) Labour market adjustments in Europe. Edward Elgar Publishing, pp 107–136. https://doi.org/10.4337/9781845428969
Le Dé L, Gaillard JC, Friesen W, Pupualii M, Brown C, Aupito A (2016) Our family comes first: migrants' perspectives on remittances in disaster. Migr Dev 5(1):130–148
Lee T, Peri G, Viarengo M (2022) The gender aspect of migrants' assimilation in Europe. Labour Econ 78:102180
Leoncini R, Macaluso M, Polselli A (2024) Gender segregation: analysis across sectoral dominance in the UK labour market. Empir Econ 67:2289–2343
Lewis E (2011) Immigration, skill mix, and capital skill complementarity. Q J Econ 126(2):1029–1069. https://doi.org/10.1093/qje/qjr011

Marra G, Radice R (2017) GJRM: generalised joint regression modelling. R package version 0.1–2. Available on CRAN.
Marra G, Radice R (2013) A penalized likelihood estimation approach to semiparametric sample selection binary response modeling. Electron J Stat 7:1432–1455
McAuliffe M, Oucho LA (2024) World migration report 2024. International Organization for Migration (IOM), Geneva
Migration Observatory (2023) The labour market effects of immigration. University of Oxford Press, Oxford
OECD (2017) Making integration work: assessment and recognition of foreign qualifications. OECD Publishing, Paris
OECD/European Commission (2023) Third-country nationals in the European Union and European OECD countries. In: OECD (ed.) Indicators of Immigrant Integration 2023: Settling In, pp. 211–243. Paris: OECD Publishing. https://doi.org/10.1787/1d5020a6-en.
Özgüzel C and A Edo (2023), Immigration and employment dynamics in European regions, OECD Regional Development Papers, No. 52, OECD Publishing, Paris, https://doi.org/10.1787/a20155a2-en
Palencia-Esteban A (2022) Occupational segregation of female and male immigrants in the European Union: accounting for cross-country differences. Int Labour Rev 161:341–373
Palencia-Esteban A, del Río C (2024) Winners and losers from occupational segregation across Europe: the role of gender and migration status. Migr Stud 12(1):21–41. https://doi.org/10.1093/migration/mnad029
Parey M, Ruhose J, Waldinger F, Netz N (2017) The selection of high-skilled emigrants. Rev Econ Stat 99(5):776–792
Pecoraro M, Wanner P (2019) Does the recognition of foreign credentials decrease the risk for immigrants of being mismatched in education or skills? In: Steiner I, Wanner P (eds) Migrants and Expats: The Swiss Migration and Mobility Nexus. IMISCOE Research Series, Springer, pp 161–186
Peri G (2016) Immigrants, productivity, and labor markets. J Econ Perspect 30(4):3–30
Peri G, Sparber C (2009) Task specialization, immigration, and wages. Am Econ J Appl Econ 1(3):135–169
Pioré MJ (1979) Birds of passage: migrant labor and industrial societies. Cambridge University Press
Poeschel F (2020) Out there on your own: absence of the spouse and migrants' integration outcomes. MPRA Paper 98993. University Library of Munich, Germany.
Reyneri E, Fullin G (2011) Labour market penalties of new immigrants in new and old receiving west European countries. Int Migr 49:31–57
Roy AD (1951) Some thoughts in the distribution of earnings. Oxf Econ Pap 3(2):135–146
Rubin DB (2005) Causal inference using potential outcomes: design, modeling, decisions. J Am Stat Assoc 100(469):322–331
Salvatore MA, Baglio ME (2018) A systematic review of working conditions and occupational health among immigrants in Europe and Canada. BMC Public Health, 18(1), 770. https://bmcpublichealth.biomedcentral.com/articles/https://doi.org/10.1186/s12889-018-5703-3.
Sjaastad LA (1962) The costs and returns of human migration. J Polit Econ 70(4):80–93
Sklar A (1973) Random variables, joint distributions, and copulas. Kybernetika 9:449–460
Smith MD (2003) Modelling sample selection using Archimedean copulas. Econometrics J 6(1):99–123
Sparreboom T, Tarvid A (2017) Skills mismatch of natives and immigrants in Europe. International Labour Organization, Geneva
Tani M (2017) Local signals and the returns to foreign education. Econ Educ Rev 61:174–190
Trivedi P, Zimmer D (2007) Copula modeling: an introduction for practitioners. Found Trends® Econ 1(1):1–111
Verschueren H (2016) Employment and social security rights of third-country labour migrants under EU law: an incomplete patchwork of legal protection. Eur J Migr Law 18(4):373–408
Winkelmann R (2011) Copula bivariate probit models: with an application to medical expenditures. Health Econ 21:1444–1455
Wright CF, McLaughlin C (2024) Short-term fix or remedy for market failure? Immigration policy as a distinct source of skills. Ind Relat J 55(1):3–19. https://doi.org/10.1111/irj.12412
Zaiceva A (2014) The impact of aging on the scale of migration. IZA World Labor 101:1–10

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.