Perceived Discrimination in the European Labor Market: Demographic Determinants and Cross-Country Variations

Xiaojing Huang

Department of Statistical Science
University of Cape Town

March 28, 2025

Overview

1. Hypotheses
2. Response variables
3. Predictors
4. Methodology
5. Results
6. Going forward

Hypotheses

- H1: Female migrants are more likely to perceive discrimination compared to the male migrants, especially the discrimination relating to "Lack of qualification", "Lack of language skills" and "No suitable job available".
- H2: Highly educated migrants are more likely to over-report discrimination compared to migrants who are less educated.
- H3: Demographic factors such as educational attainment, country of birth, age, hosting country and gender significantly predict the likelihood of migrants perceiving discrimination in Europe.
- H4: There are distinct groups of European countries where the migrants share simliar discrimination perception based on the demographical backgrounds of migrants in those countries.

Response variables

Response_variable	Total_observations_in_thousands
Citizenship or residence permit	356.6
Discrimination due to foreign origin	500.1
Lack of language skills	2363.5
Lack of recognition of qualifications	1346.0
Language skills, qualifications, citizenship, foreign origin, job and other barriers	15959.9
Never sought work or never worked	9270.6
No suitable job available	1263.0
None	44268.0
Other	4370.2

Predictors

Predictor	Levels
Age	15-24 years old
	25-54 years old
	55-74 years old
Gender	Female
	Male
Educational_attainment	Less than primary, primary and lower secondary education (Level 0-2)
	Tertiary education (Level 5-8)
	Upper secondary and post-secondary non-tertiary education
Country_of_birth	EU27 except reporting country
	Foreign country
	Non-EU27 countries (from 2020) nor reporting country
Hosting_country	Austria, Belgium, Croatia, Cyprus, Czechia
	Denmark, Estonia, Finland, France, Germany
	Greece, Hungary, Ireland, Italy, Latvia
	Lithuania, Luxembourg, Malta, Netherlands, Norway
	Poland, Portugal, Slovakia, Slovenia, Spain
	Sweden, Switzerland

Methodology

1. Multinomial Logistic Regression Model:
   - Using Cramér's V to check the correlation between the predictors.
   - Fitting the regression model with necessary interaction terms.
   - Assesing the effect of Gender and Educational attainment (together with the interaction terms) on the likelihood of migrants perceiving discrimination to test H1 and H2.
   - Testing the predictive power of these variables by using the Fisher exact test and also AIC/BIC scores to test H3.
2. Data transformation, PCA and K-means clustering
   - Transform dataset from categorical to numeric by calculating for each hosting country the proportion of observations allocated to each level of each predictor.
   - Use log-ratio to transform the dataset again and perform PCA to find uncorrelated variables.
   - Based on the result of PCA perform K-means with optimal number of clusters (Using Elbow Method).

Results - Cramér's V

	Country_of_birth	Educational_attainment	Age	Sex	Hosting_country
Country_of_birth	1.0000000	0.0371805	0.0233559	0.0363961	0.1326292
Educational_attainment	0.0371805	1.0000000	0.1606308	0.0308024	0.1326317
Age	0.0233559	0.1606308	1.0000000	0.0504320	0.1541855
Sex	0.0363961	0.0308024	0.0504320	1.0000000	0.0678646
Hosting_country	0.1326292	0.1326317	0.1541855	0.0678646	1.0000000

Results - Hypothesis 1

H1: Female migrants are more likely to perceive discrimination compared to the male migrants, especially the discrimination relating to “Lack of qualification”, “Lack of language skills” and “No suitable job available”.

y.level	term	estimate	std.error	statistic	p.value	odds_ratio
Citizenship or residence permit	SexMales	0.0655570	0.0895751	0.7318664	0.4642501	1.0677536
Discrimination due to foreign origin	SexMales	-0.3594915	0.0808392	-4.4469958	0.0000087	0.6980312
Lack of language skills	SexMales	-0.4293914	0.0432072	-9.9379553	0.0000000	0.6509051
Lack of recognition of qualifications	SexMales	-0.7536204	0.0576410	-13.0743823	0.0000000	0.4706595
Language skills, qualifications, citizenship, foreign origin, job and other barriers	SexMales	-0.3558047	0.0195262	-18.2219501	0.0000000	0.7006094
Never sought work or never worked	SexMales	-1.4154465	0.0291794	-48.5084710	0.0000000	0.2428172
No suitable job available	SexMales	-0.5390726	0.0575099	-9.3735616	0.0000000	0.5832890
Other	SexMales	-0.2945810	0.0338589	-8.7002450	0.0000000	0.7448436

Results - Hypothesis 2

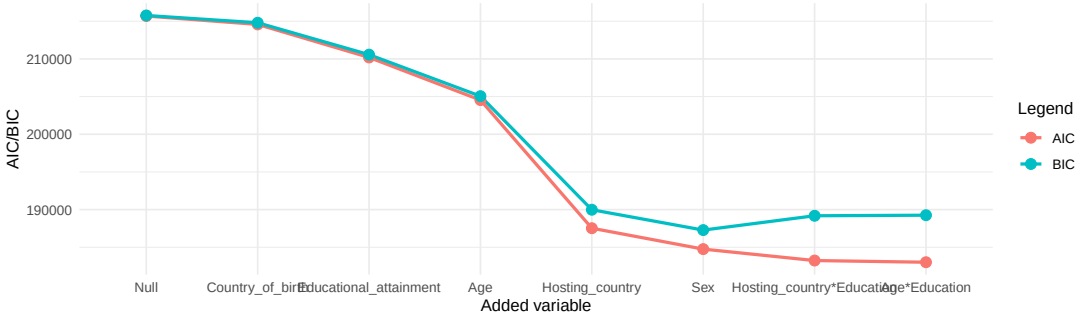
H2: Highly educated migrants are more likely to over-report discrimination compared to migrants who are less educated.

y.level	term	estimate	std.error	statistic	p.value	odds_ratio
Citizenship or residence permit	Educational_attainmentTertiary education (levels 5-8)	-1.5216114	1.7947412	-0.8478166	0.3965401	0.2183597
Discrimination due to foreign origin	Educational_attainmentTertiary education (levels 5-8)	-2.1351583	2.2065849	-0.9676303	0.3332291	0.1182259
Lack of language skills	Educational_attainmentTertiary education (levels 5-8)	-0.4545019	0.6165285	-0.7371952	0.4610036	0.6347641
Lack of recognition of qualifications	Educational_attainmentTertiary education (levels 5-8)	-2.6730556	4.0555060	-0.6591177	0.5098202	0.0690409
Language skills, qualifications, citizenship, foreign origin, job and other barriers	Educational_attainmentTertiary education (levels 5-8)	-0.2822824	0.2542471	-1.1102678	0.2668836	0.7540607
Never sought work or never worked	Educational_attainmentTertiary education (levels 5-8)	-4.2199249	0.3983035	-10.5947466	0.0000000	0.0146997
No suitable job available	Educational_attainmentTertiary education (levels 5-8)	-1.9834970	1.8230383	-1.0880172	0.2765875	0.1375872
Other	Educational_attainmentTertiary education (levels 5-8)	-3.1006747	1.4822578	-2.0918592	0.0364511	0.0450188

Results - Hypothesis 2

y.level	term	estimate	std.error	statistic	p.value	odds_ratio
Language skills, qualifications, citizenship, foreign origin, job and other barriers	Educational_attainmentTertiary education (levels 5-8):AgeFrom 25 to 54 years	1.0736839	0.2184530	4.914942	0.0000009	2.926139
Language skills, qualifications, citizenship, foreign origin, job and other barriers	Educational_attainmentTertiary education (levels 5-8):AgeFrom 55 to 74 years	1.3077426	0.2238707	5.841508	0.0000000	3.697817
Never sought work or never worked	Educational_attainmentTertiary education (levels 5-8):AgeFrom 25 to 54 years	3.6645967	0.3482606	10.522571	0.0000000	39.040386
Never sought work or never worked	Educational_attainmentTertiary education (levels 5-8):AgeFrom 55 to 74 years	3.6332845	0.3503233	10.371233	0.0000000	37.836886
Never sought work or never worked	Educational_attainmentTertiary education (levels 5-8):Hosting_countrySpain	0.8730675	0.2154782	4.051767	0.0000508	2.394244
Other	Educational_attainmentTertiary education (levels 5-8):AgeFrom 25 to 54 years	3.4340825	1.1411907	3.009210	0.0026193	31.002953

Results - Hypothesis 3



Results - Hypothesis 4

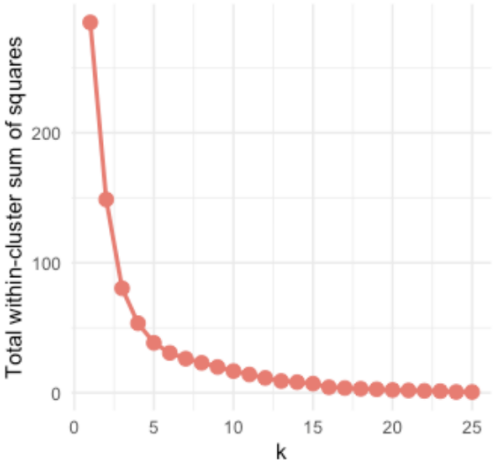
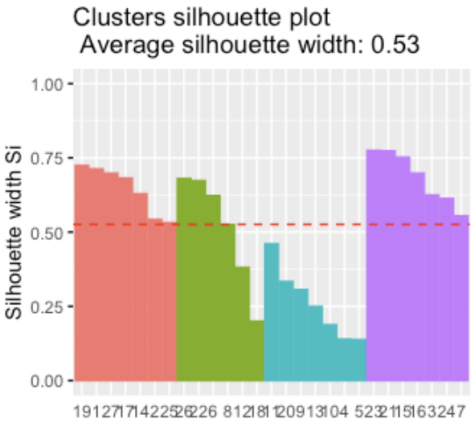


Figure2: Change in total within cluster sum of squares



Silhouette Analysis for K-Means Clustering

Results - Hypothesis 4

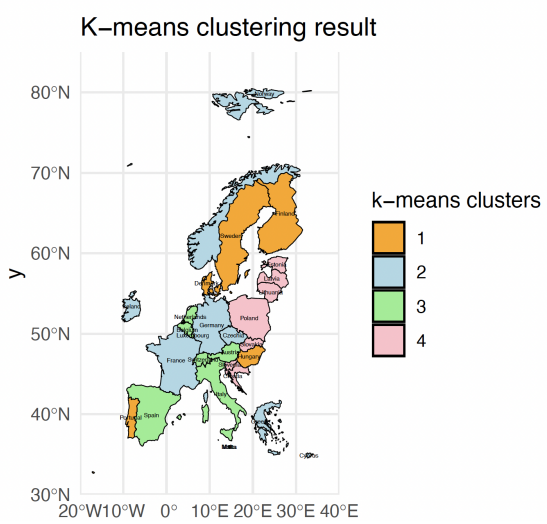


Table: Assignment of countries to the clusters

cluster	countries
1	Austria, Belgium, Italy, Luxembourg, Netherlands, Spain, Switzerland
2	Denmark, Finland, Hungary, Malta, Portugal, Sweden
3	Cyprus, Czechia, France, Germany, Greece, Ireland, Norway
4	Croatia, Estonia, Latvia, Lithuania, Poland, Slovakia, Slovenia

Going forward

- Due to the transformation significantly reduced the dataset size, different approaches such as using Latent class analysis would be indirectly test the hypothesis by uncovering the latent subgroups of migrants based on their demographic backgrounds and perceived discrimination at individual-level instead of country-level.
- After identifying the latent classes, the study will aim to further run a structural equations model in order to examine whether the structural relationship between variables differs across the latent classes to see how potential policy recommendations can address the discrimination accordingly.