

Counting Shadows: Estimating Unauthorised Migrant Populations

Maciej Beręsewicz^{1,2} **Brendan Georgeson³** **Peter Walsh⁴**
Aniela Czerniawska⁵ **Li-Chun Zhang^{6,7}**

¹⁾ Poznań University of Economics and Business, Poland, ²⁾ Statistical Office in Poznań, Poland ³⁾ Office for National Statistics ⁴⁾ Migration Observatory, University of Oxford ⁵⁾ Adam Mickiewicz University, Poland ⁶⁾ University of Southampton ⁷⁾ Statistisk sentralbyrå (Statistics Norway)

COMPAS Work in Progress Series, University of Oxford, 13.11.2025



POZNAŃ UNIVERSITY
OF ECONOMICS
AND BUSINESS

Contents

1 Introduction

2 Motivation

3 Theoretical model

4 Empirical model

5 Selected results for 18+ population

6 Conclusions

7 Literature (selected)

Introduction

- In this paper, we estimate the size of the unauthorised (non-Schengen 18+) foreign-born population in Poland.
- We propose a theoretical and empirical model that uses only aggregated data on the observed unauthorised, lawful, and criminal populations based on Zhang (2008) work.
- We consider the size of the unauthorised population as random which fundamentally differs from the *residual method* (Passel et al. 2004):
 - ① We treat the unauthorised population M_{it} as an inherently random variable rather than an unknown constant.
 - ② We explicitly model the structural relationship between authorised and unauthorised populations through $\xi_{it} = \mathbb{E}(M_{it}|N_{it})$, capturing how these populations co-evolve and relate through migrant networks.
 - ③ Relating ξ_{it} to the regular population N_{it} helps to control for extraneous variation (e.g., seasonality, economic cycles, or demographic trends) by normalizing the unauthorised population relative to a stable reference frame.

Introduction

- Based on the data from 2019 to 2023 for Poland, we conclude that:
 - the size of the unauthorised population in Poland varied from 40,000 to 150,000 depending on the parameterisation (i.e., time varying, sex varying coefficients),
 - its relation to the reference (authorised) population varied between 2% and 10% (with an exception for 2019),
 - it does not matter what reference population we use: Population, Tax or Social Insurance register,
 - its relation to the whole regular Poland population (about 38 millions) is negligible (less than 0.5%).
- Finally, we discuss possibilities of applying this method to the UK.

Contents

1 Introduction

2 Motivation

3 Theoretical model

4 Empirical model

5 Selected results for 18+ population

6 Conclusions

7 Literature (selected)

Poland as a case study – borders

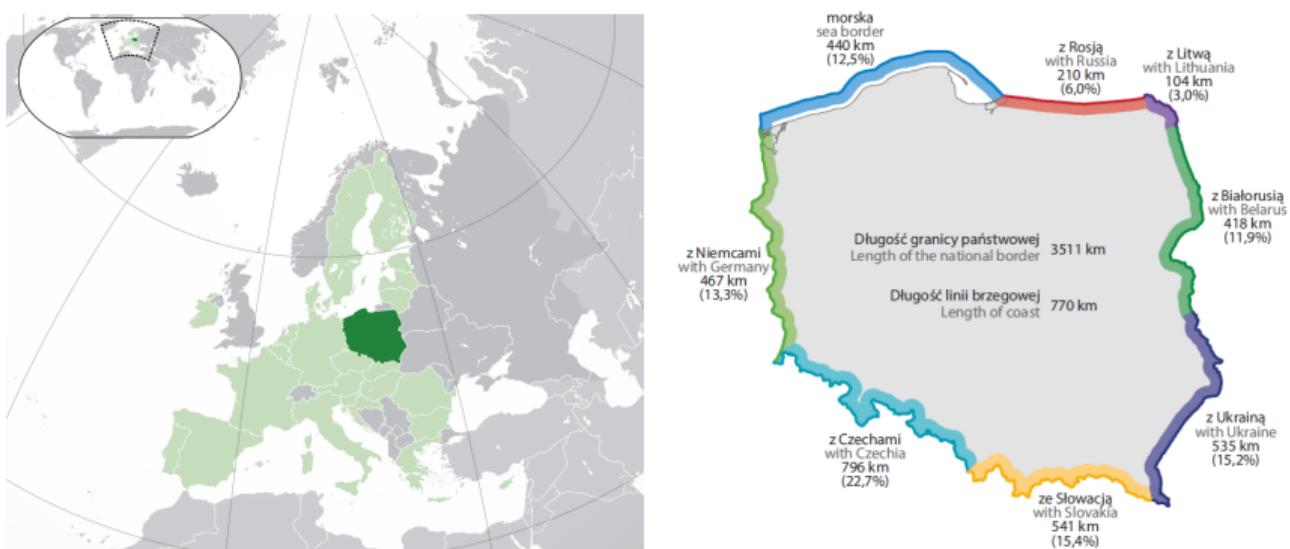


Figure 1: Poland and its borders

Poland as a case study – migration

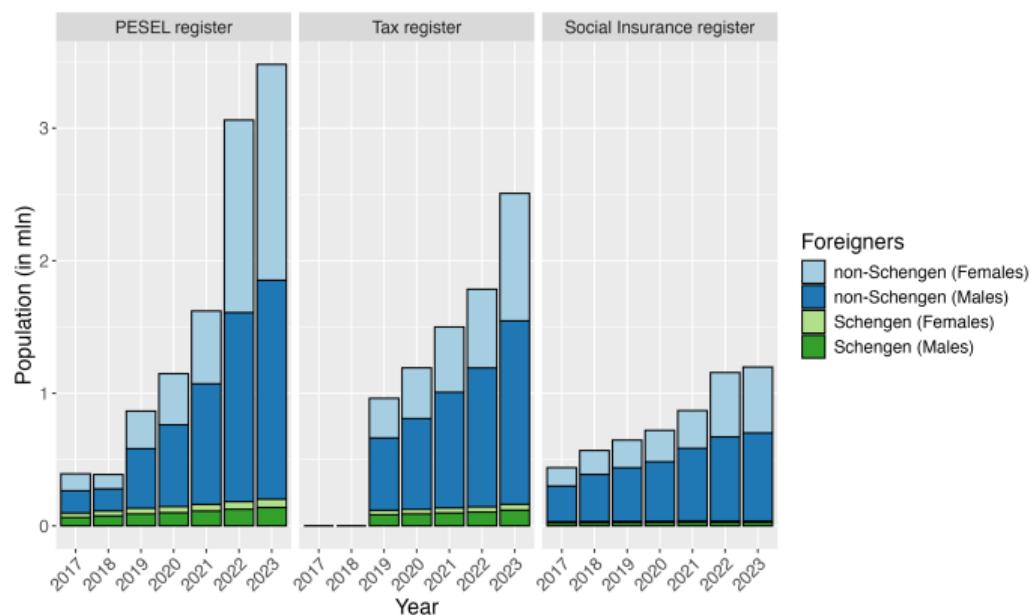


Figure 2: Migration to Poland between 2017 and 2023 measured by three registers: Population register (PESEL), Tax and Social Insurance (employees/employed) register

Poland as a case study – ongoing border crisis

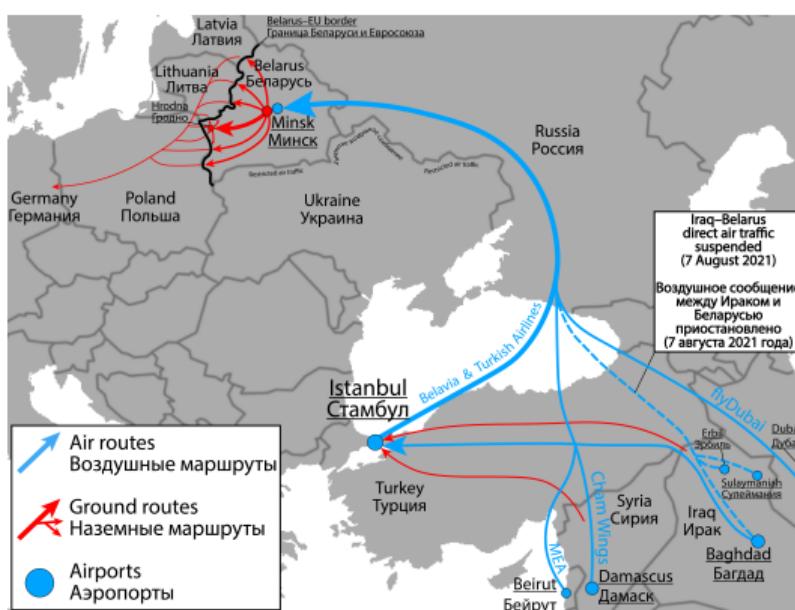


Figure 3: Belarus–European Union border crisis (from August 2021). Source:
https://en.wikipedia.org/wiki/Belarus%E2%80%93European_Union_border_crisis

Poland as a case study – unauthorised migration

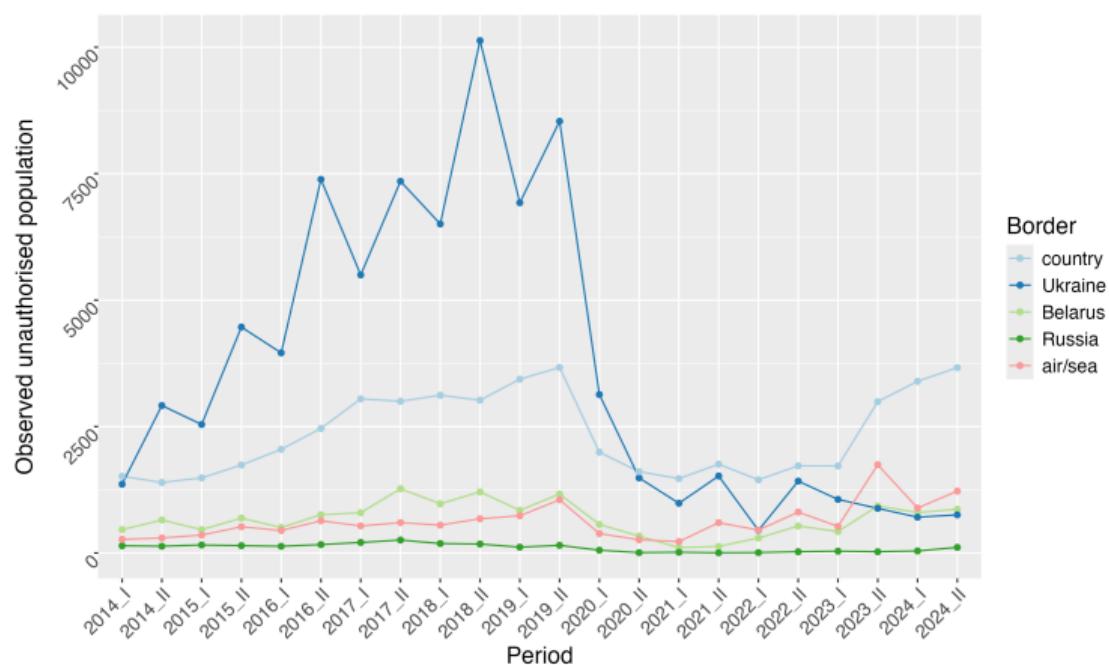


Figure 4: Unauthorised migration in Poland by the place of apprehension between 2014 and 2024

Poland as a case study – unauthorised migration within country

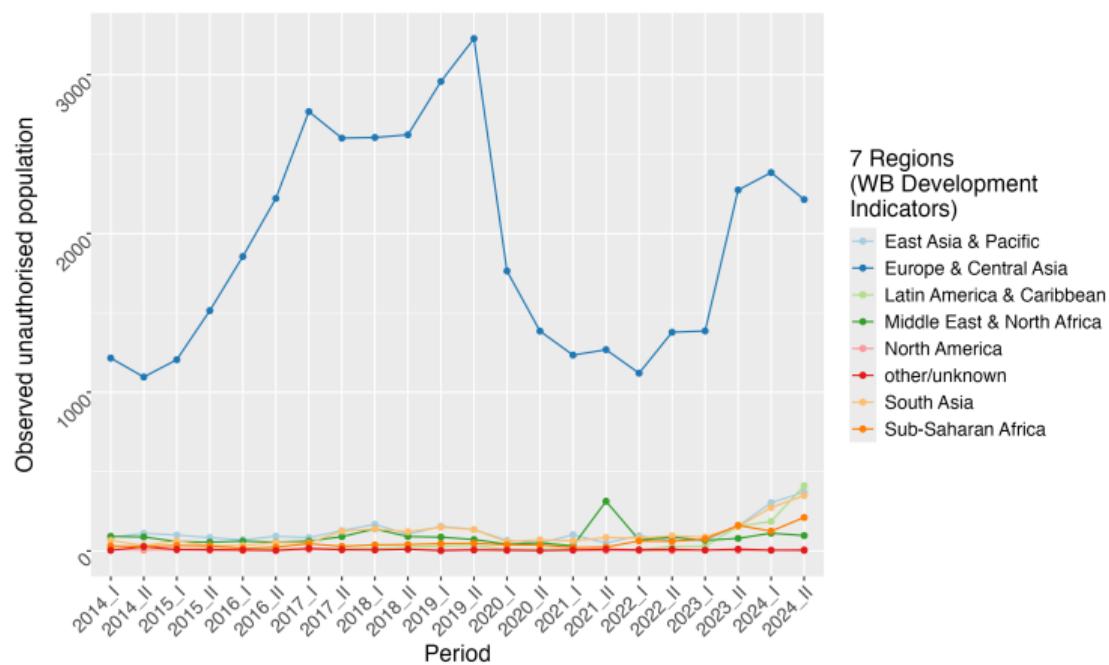


Figure 5: Unauthorised migration within Poland between 2014 and 2024 by regions defined by the World Bank

Contents

1 Introduction

2 Motivation

3 Theoretical model

4 Empirical model

5 Selected results for 18+ population

6 Conclusions

7 Literature (selected)

Theoretical model – notation

- Let M_{it} be the size of the *unauthorised* population for the i -th country ($i = 1, \dots, I$) at period t ($t = 1, \dots, T$).
- Let N_{it} be the size of the *reference* (e.g., authorized) population for the same country at the same period.
- Let m_{it} be the *observed* counts of the unauthorised population (being a part of M_{it}).
- Let n_{it} be the observed counts of the (possibly) authorized population that serves the role of a covariate that can explain the levels of m_{it} .
- For each country, M_{it} is unobserved and the triplet (N_{it}, m_{it}, n_{it}) is observed.

Theoretical model – theoretical size of the M_{it}

- We focus on the *theoretical* size of the M_{it} which is defined as the following conditional expectation

$$\xi_{it} = \mathbb{E}(M_{it}|N_{it}). \quad (1)$$

- In addition, we define p_{it} as the *detection rate* of the m_{it} and the ρ_{it} as the *theoretical* detection rate of the observed unauthorised population which is given by the following conditional expectation

$$\rho_{it} = \mathbb{E}(p_{it}|N_{it}, n_{it}). \quad (2)$$

- Further, we can combine (1) and (2) into the expected counts as

$$\mathbb{E}(m_{it}) = \mathbb{E}(M_{it}|N_{it})\mathbb{E}(p_{it}|N_{it}, n_{it}). \quad (3)$$

- We assume that (1) and (2) are conditionally independent given the covariates N_{it} and n_{it} .

Theoretical model – evidence from the migration literature

- Empirical studies consistently show that unauthorised migrants depend **heavily on co-ethnic and kin-based networks** for survival and incorporation into host societies (Hagan, 1998; Engbersen et al., 2000; Ambrosini & Hager, 2023).
- Such networks play multiple roles – *offering accommodation, information, informal job placements, and emotional support* – and often function as substitutes for the formal welfare and employment systems that unauthorised migrants cannot access.
- In the United Kingdom, qualitative studies of cleaning and domestic work reveal how immigrant networks, multi-layered subcontracting, and employer practices create an “invisible” niche characterised by low pay, precarity, and limited institutional protection (Anderson, 2007; McDowell, Batnitzky, & Dyer, 2009).
- Ambrosini and Hager (2023) argue that co-ethnic **networks are critical mediators** between local labour demand and unauthorised migrant supply.

Contents

1 Introduction

2 Motivation

3 Theoretical model

4 Empirical model

5 Selected results for 18+ population

6 Conclusions

7 Literature (selected)

Empirical model – models for ξ and ρ

- For ξ_{it} we assume

$$\xi_{it} = \mathbb{E}(M_{it}|N_{it}) = N_{it}^{\alpha}, \quad (4)$$

where α is unknown and should be estimated from the data. If we have enough information we can postulate country specific parameter α_i or country-period specific α_{it} .

- Furthermore, for ρ_{it} we assume

$$\rho_{it} = \mathbb{E}(p_{it}|N_{it}, n_{it}) = \left(\frac{n_{it}}{N_{it}} \right)^{\beta}, \quad (5)$$

where β is unknown and should be estimated from the data.

- Moreover, if we assume that α and β could be supplemented with additional covariates (e.g. year effect) we can include additional covariates, say z_{it} .

Empirical model – the count regression

- Now, to estimate α and β we need to postulate the model for the observed counts m_{it} . The natural selection would be Poisson distribution,

$$m_{it} \sim \text{Poisson}(\lambda_{it}), \quad (6)$$

where $\lambda_{it} = \mathbb{E}(M_{it}|N_{it})\mathbb{E}(p_{it}|N_{it}, n_{it})$ which translates into the following empirical model under (1) and (2)

$$m_{it} \sim \text{Poisson} \left\{ N_{it}^{\alpha} \left(\frac{n_{it}}{N_{it}} \right)^{\beta} \right\}. \quad (7)$$

- In the original publication Zhang (2008) used Negative Binomial model, but we may consider another as long as the expected is the same as in above.
- We can also consider OLS: $\log(m_{it}) = \alpha \log(N_{it}) + \beta \log(n_{it}/N_{it})$ based on carefully aggregated data to avoid 0s in m_{it} and n_{it} .

Empirical model – estimator for the unauthorised population size

- After we obtain α and β (or its variants for country and period) we can estimate the size of the unauthorised population given by

$$\hat{\xi}_t = \sum_{i=1}^I N_{it}^{\hat{\alpha}} \quad \text{or} \quad \hat{\xi}_t = \sum_{i=1}^I N_{it}^{\hat{\alpha}_{it}}, \quad (8)$$

- Due to non-linear character the proposed estimator given in (8) will be biased thus we may consider using a bias corrected estimator given by

$$\hat{\xi}_t^{bc} = \hat{\xi}_t - \frac{1}{2} \left(\sum_{i=1}^I N_{it}^{\hat{\alpha}} [\ln(N_{it})]^2 \right) \widehat{\text{Var}}[\hat{\alpha}], \quad (9)$$

where we plug-in estimated α and its variance.

- In the paper we prove that the (8) is consistent and derive the bias-corrected estimator.
- In this presentation we report uncorrected estimates only.

Contents

1 Introduction

2 Motivation

3 Theoretical model

4 Empirical model

5 Selected results for 18+ population

6 Conclusions

7 Literature (selected)

Outline

- 5 Selected results for 18+ population
 - Poland
 - The UK

Poland – data (non-Schengen)

Table 1: Number of non-Schengen foreigners observed in three registers: Border Guards, Police, Prison, and three reference population registers: PESEL, Tax, Employees

Year	Sex	Border (<i>m</i>)	Auxiliary (<i>n</i>)		Reference population (<i>N</i>)		
			Police	Prison	PESEL	Tax	Employees
2019	F	1,538	734	50	283.6	299.4	208.8
	M	5,088	9,154	1,000	449.9	544.8	403.4
2020	F	699	687	49	385.6	382.1	236.8
	M	2,706	8,797	973	617.2	684.5	448.1
2021	F	485	935	49	549.5	492.2	284.7
	M	2,625	12,012	1,434	908.7	872.3	546.8
2022	F	319	1,399	47	1,451.9	594.2	485.7
	M	2,642	15,363	1,520	1,426.3	1,047.0	633.6
2023	F	525	1,791	82	1,628.5	961.8	497.4
	M	3,848	16,253	1,988	1,651.0	1,382.7	663.4

Data and distributional assumptions

- Following Zhang (2008), we have aggregated data due to a large number of zeros in m and n . Aggregation was conducted as follows:
 - we check which countries meet the criteria: $m_{it} > 0$, $n_{it} > 0$, $N_{it} > 0$, and $n_{it}/N_{it} > 0$,
 - all countries that do not meet these criteria were combined into one called rest,
 - we have conducted the same procedure for sex as a covariate for α and β .
- We have visually verified whether the assumptions of the model are met, i.e., negative correlation between $\log(N_{it})$ and $\log(m_{it}/N_{it})$ and positive correlation between $\log(n_{it}/N_{it})$ and $\log(m_{it}/N_{it})$.
- In this study, we considered three models (OLS on $\log(m_{it})$, Poisson and Negative Binomial on m_{it}) without and with *time trend* and *sex* as covariates.
- Variance was calculated using fractional weighted bootstrap.
- We have conducted robustness checks, outlier-robust alternatives, and quality assessment of these models.

Assumptions – correlations

Table 2: Correlation between reference population (N), auxiliary covariate (n), and observed unauthorised population (m) by reference population, overall or with sex between 2019 and 2023

Reference Population	Type	$\log(N), \log(m/N)$		$\log(n/N), \log(m/N)$	
		Pearson	Spearman	Pearson	Spearman
PESEL	Overall	-0.607	-0.611	0.713	0.715
	With Sex	-0.642	-0.634	0.733	0.735
Tax	Overall	-0.610	-0.605	0.723	0.691
	With Sex	-0.645	-0.637	0.730	0.722
Employees	Overall	-0.564	-0.516	0.631	0.553
	With Sex	-0.606	-0.564	0.659	0.613

Note: Pearson correlation measures a linear relationship, while Spearman correlation measures a monotonic (non-linear but consistent) relationship. Spearman is based on ranks instead of original values and is robust to outliers.

Assumptions – visual inspection (Tax register, overall)

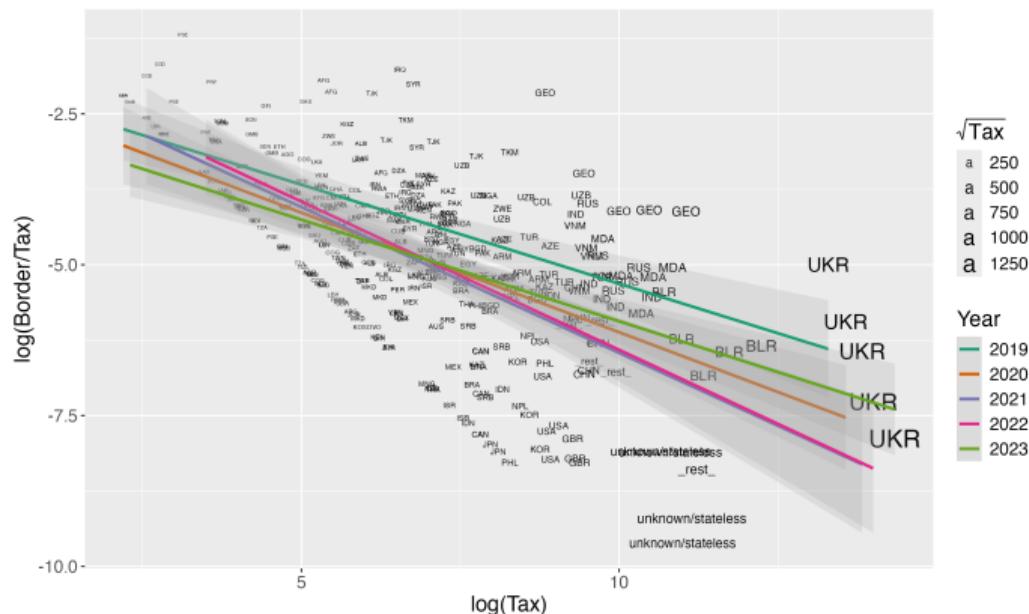


Figure 6: Relation between $\log(N)$ and $\log(m/N)$ between 2019 and 2023

Assumptions – visual inspection (Tax register, overall)

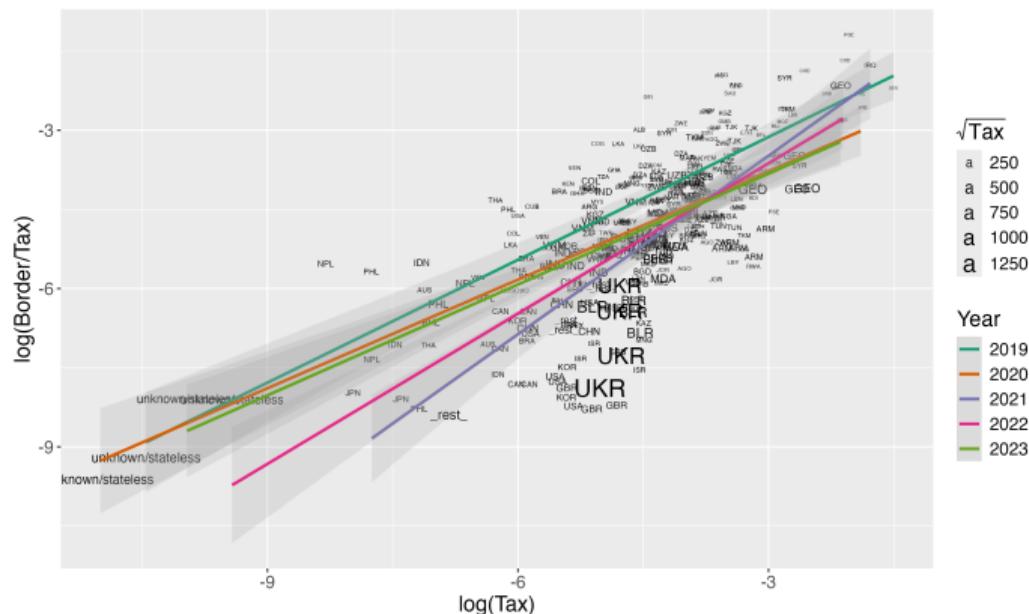


Figure 7: Relation between $\log(n/N)$ and $\log(m/N)$ between 2019 and 2023

Estimated relation of the unauthorised to the authorised population(s)

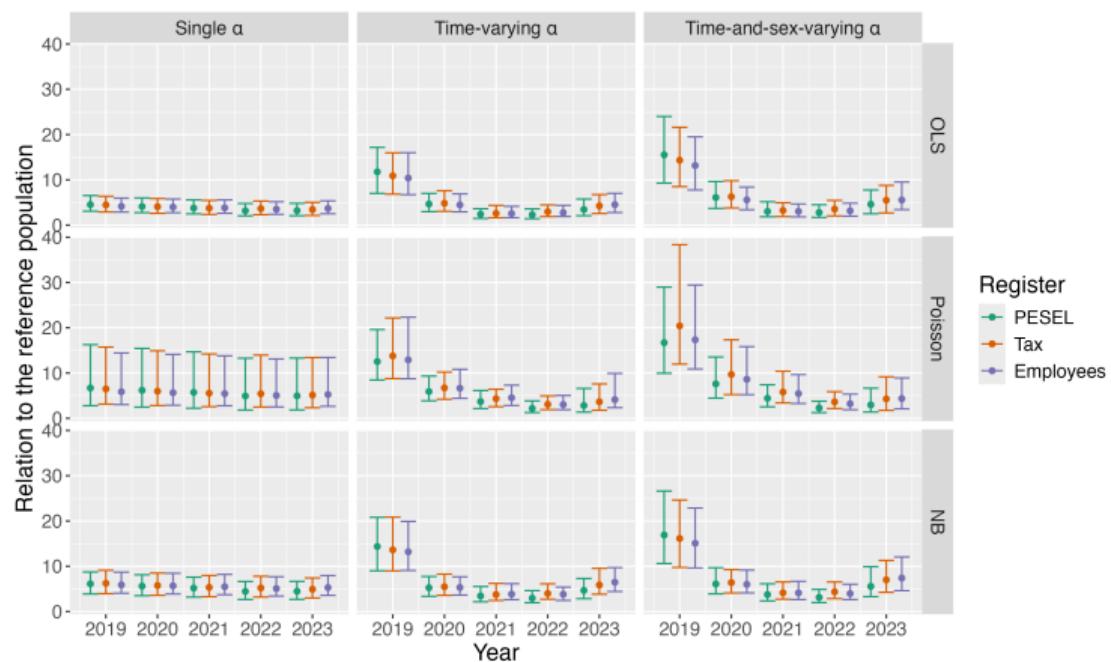


Figure 8: Point estimates and 95% confidence intervals for the ratio to the reference population by model and covariates

Outline

- 5 Selected results for 18+ population
 - Poland
 - The UK

The UK

Table 3: Basic information about unauthorised migration to the UK measured by enforced and voluntary returns, population size, and prison population between 2021 and 2024

Year	Return		Population	Prison
	Enforced	Voluntary		
2021	2,786	7,277	8,709,207	9,744
2022	3,800	10,920	8,799,791	9,615
2023	6,344	20,270	8,923,387	10,240

The UK – assumptions

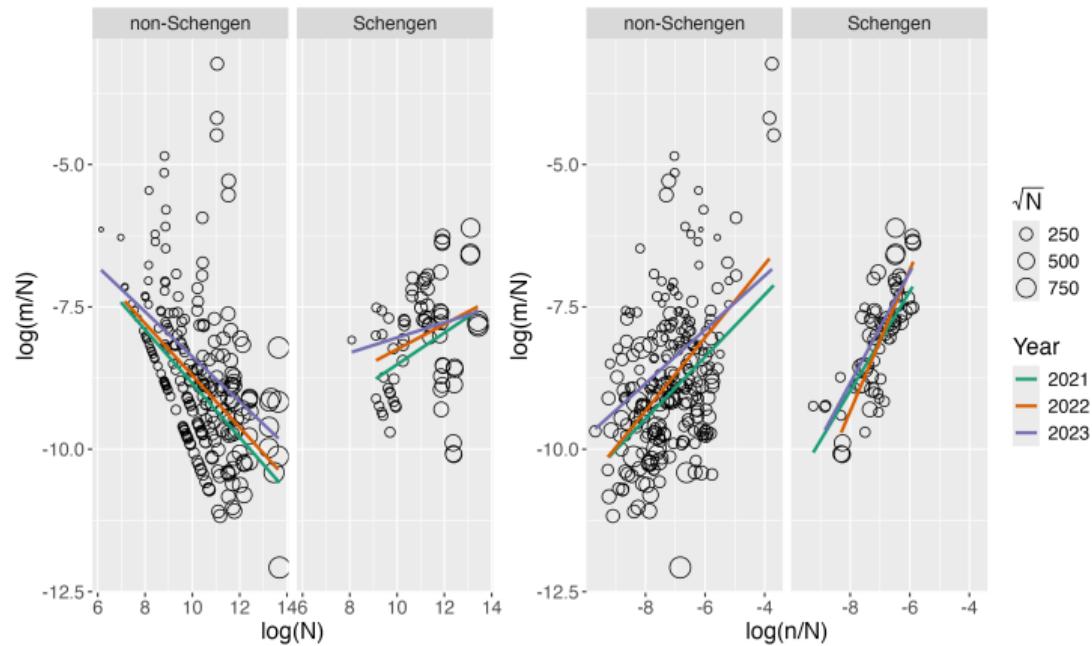


Figure 9: Verification of assumptions using enforced return as m and prison records as n

The UK – assumptions

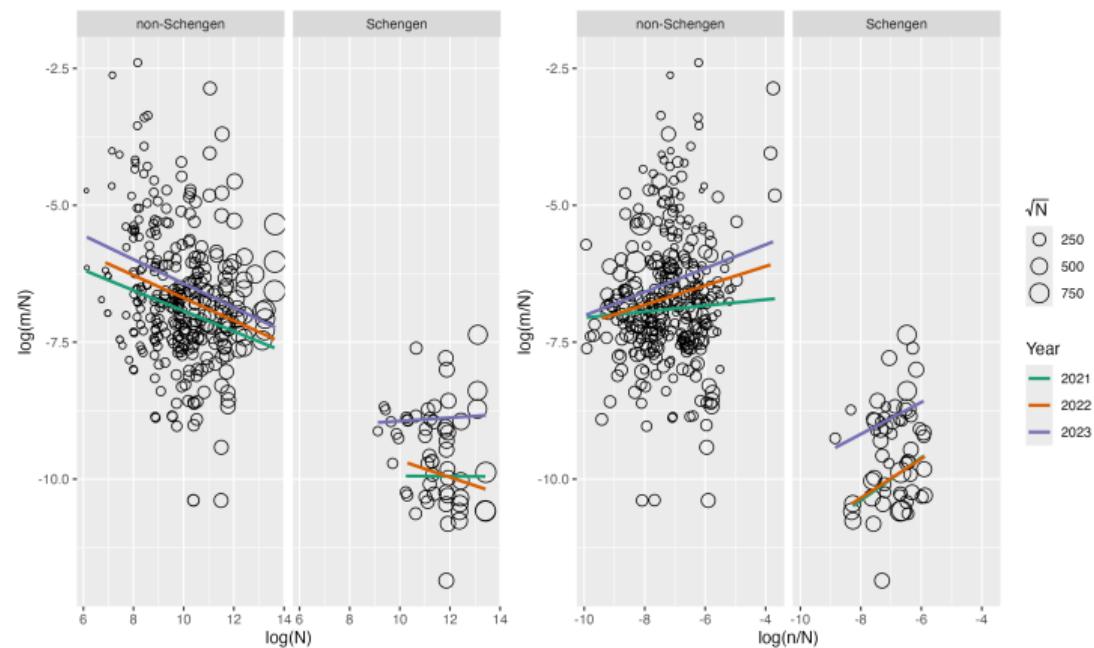


Figure 10: Verification of assumptions using voluntary return as m and prison records as n

The UK - summary

- The data analysis confirms that the assumptions hold for the UK non-Schengen foreign-born population.
- Unlike residual methods that require detailed demographic breakdowns, this model can operate with limited but structured data.
- In the UK, even basic data like returns and prison population can be used to generate initial estimates.
- This model provides a structured, evidence-based approach to support policy planning, resource allocation, and public discourse with more grounded estimates.
- Currently the possibilities of application in the UK are limited due to access of auxiliary information, that is police records by nationality. However the Office for National Statistics will progress this work in 2026 with the ultimate aim of producing a national estimate.

Contents

1 Introduction

2 Motivation

3 Theoretical model

4 Empirical model

5 Selected results for 18+ population

6 Conclusions

7 Literature (selected)

Conclusions

- We have extended Zhang (2008) model for different parametrisation and distributional assumptions. Some of the works were made earlier in the publication Beręsewicz & Pawlukiewicz (2020).
- Empirical results suggest that assumptions of the model hold for Poland and the UK.
- The size of the unauthorised population is negligible to the Population of Poland.
- Further works should focus on inclusion of children (below 18) population as well as inclusion of country-specific coefficients that captures the migration strategies (e.g. different rules of entry for the labour market).

Acknowledgements

- We would like to thank representatives of Polish Border Guards and Polish Police for making the data available.
- Maciej Beręsewicz and Aniela Czerniawska acknowledge support by the National Science Centre in Poland OPUS 20 grant no. 2020/39/B/HS4/00941.

Contents

1 Introduction

2 Motivation

3 Theoretical model

4 Empirical model

5 Selected results for 18+ population

6 Conclusions

7 Literature (selected)

Literature (selected)

- Ambrosini, M., & Hager, M. H. J. (2023). Defining and explaining irregular migration. In Irregular Migration (IMISCOE Research Series). Springer. <https://doi.org/10.1007/978-3-031-30838-3>
- Anderson, B. (2007). A very private business: Exploring the demand for migrant domestic workers. European Journal of Women's Studies, 14(3), 247–264. <https://doi.org/10.1177/1350506807079013>
- Beręsewicz, M., & Pawlukiewicz, K. (2020). Estimation of the number of irregular foreigners in Poland using non-linear count regression models. arXiv preprint arXiv:2008.09407.
- Engbersen, G., van der Leun, J., & Staring, R. (2000). How illegal immigrants settle and survive: The social embeddedness of illegal immigrants in ethnic communities. [Working paper]. Erasmus University Rotterdam.
- Hagan, J. M. (1998). Social networks, gender, and immigrant incorporation: Resources and constraints. American Sociological Review, 63(1), 55–67. <https://doi.org/10.2307/2657477>
- McDowell, L., Batnitzky, A., & Dyer, S. (2009). Precarious work and economic migration: Emerging immigrant divisions of labour in Greater London's service sector. International Journal of Urban and Regional Research, 33(1), 3–25. <https://doi.org/10.1111/j.1468-2427.2009.00831.x>
- Passel, Jeffrey S., Jennifer Van Hook & Frank D. Bean. 2004. "Estimates of the Legal and Unauthorized Foreign-born Population for the United States and Selected States, Based on Census 2000." Washington, D.C.: U.S. Census Bureau, June.
- Zhang, L. C. (2008). Developing methods for determining the number of unauthorized foreigners in Norway.

8 Appendix

8

Appendix

- Migration to Poland
- More about the data
- Assumptions – registers
- Assumptions – registers by continent
- Detailed results
- Sensitivity analysis

Poland: definitions

- Illegal stay (Poland) – stay not in accordance with the regulations concerning the conditions of entry of foreigners into the territory of the Republic of Poland and their stay in that territory.
- Illegal stay on the territory of the Republic of Poland occurs in particular when a foreigner:
 - Does not have a valid visa or other valid document entitling him/her to enter and stay on the territory of the Republic of Poland.
 - Has not left the territory of the Republic of Poland after the expiry of the permitted period of stay.
 - Has illegally crossed or attempted to illegally cross the border.
 - Is or has been working illegally.
 - Has undertaken economic activity in violation of the applicable regulations.
 - Does not have sufficient financial means to stay in the territory of the Republic of Poland.
 - Has been flagged for refusal of entry in the SIS or national list.

Migration to Poland – regulations (selected)

- Special treatment for the following countries: Armenia, Azerbaijan, Belarus, Georgia, Moldova and Ukraine. In addition, Russia (Kaliningrad Oblast)
- Ukraine: after the full scale war (The UKR status, refugees).

Poland as a case study – unauthorised migration

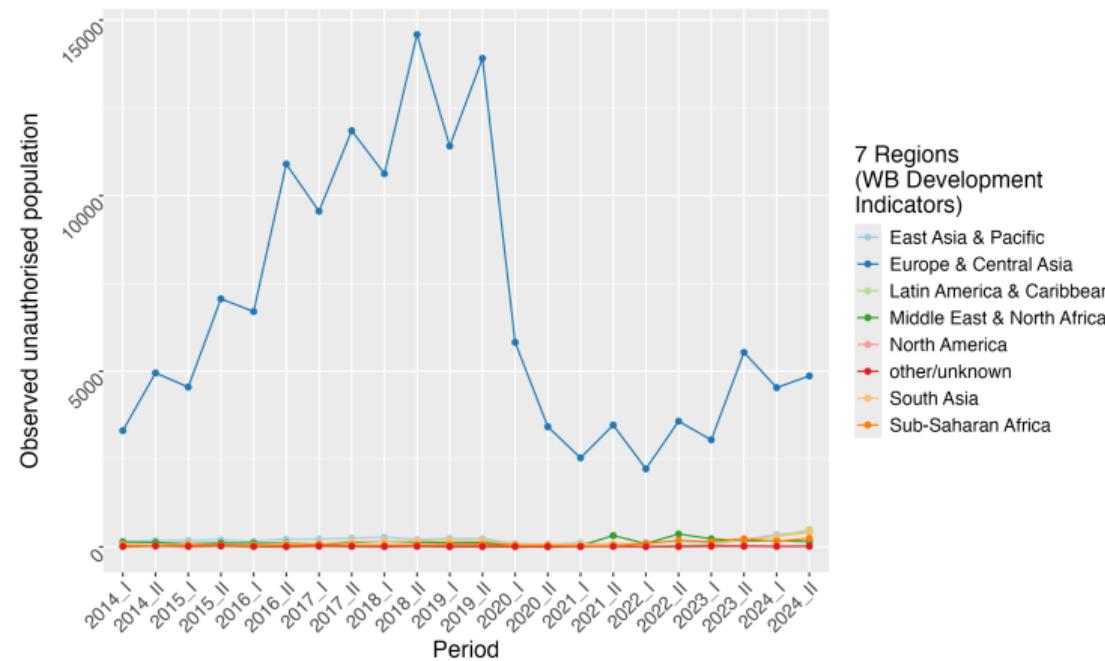


Figure 11: Unauthorised migration in Poland by the place of apprehension between 2014 and 2024 by regions defined by the World Bank

8 Appendix

- Migration to Poland
- More about the data
- Assumptions – registers
- Assumptions – registers by continent
- Detailed results
- Sensitivity analysis

Number of countries

Table 4: Number of countries by data source

Year	Border	Police	Prison	PESEL	Tax	Employees
2019	89	87	51	158	159	112
2020	76	86	53	159	161	115
2021	64	92	51	161	160	120
2022	75	96	51	163	163	128
2023	97	105	53	163	162	129

Number of countries meeting criteria

Criterion: $N_{it} > 0$, $n_{it} > 0$, $m_{it} > 0$ and $n_{it}/N_{it} < 1$ for a given reference population.

Table 5: Number of countries by the reference population

Year	PESEL	TAX	Employees
2019	74	72	69
2020	60	59	58
2021	59	59	58
2022	70	70	68
2023	86	85	83

The *rest* group by reference population

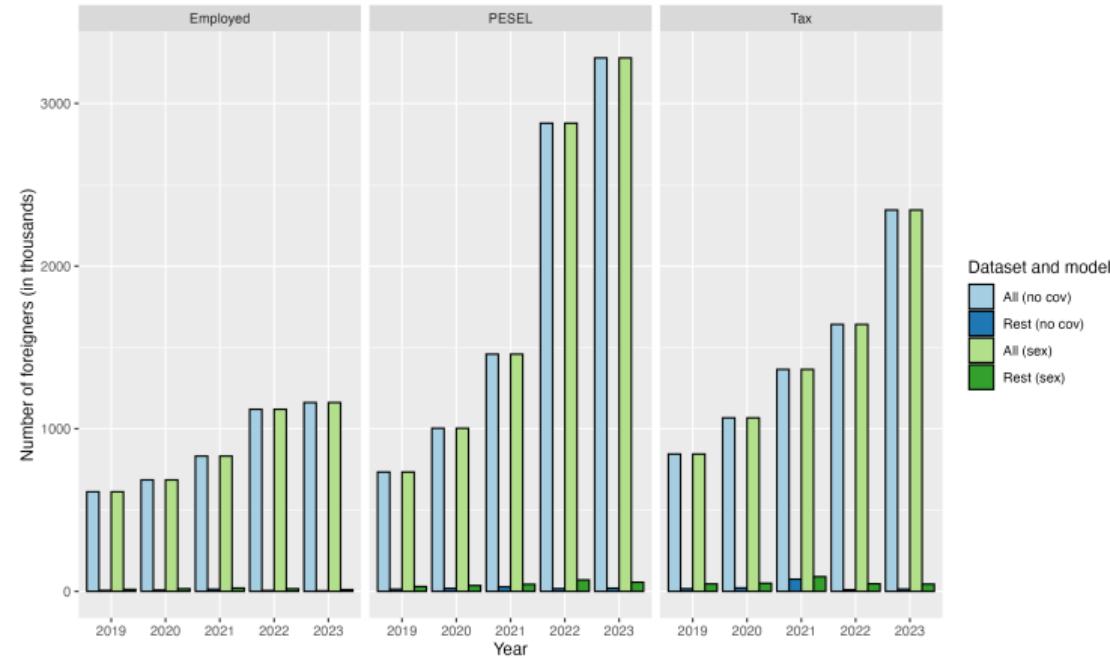


Figure 12: The *rest* group by reference population for the period 2019 to 2023

8 Appendix

- Migration to Poland
- More about the data
- **Assumptions – registers**
 - Assumptions – registers by continent
 - Detailed results
 - Sensitivity analysis

Assumptions – visual inspection (Tax register, by sex)

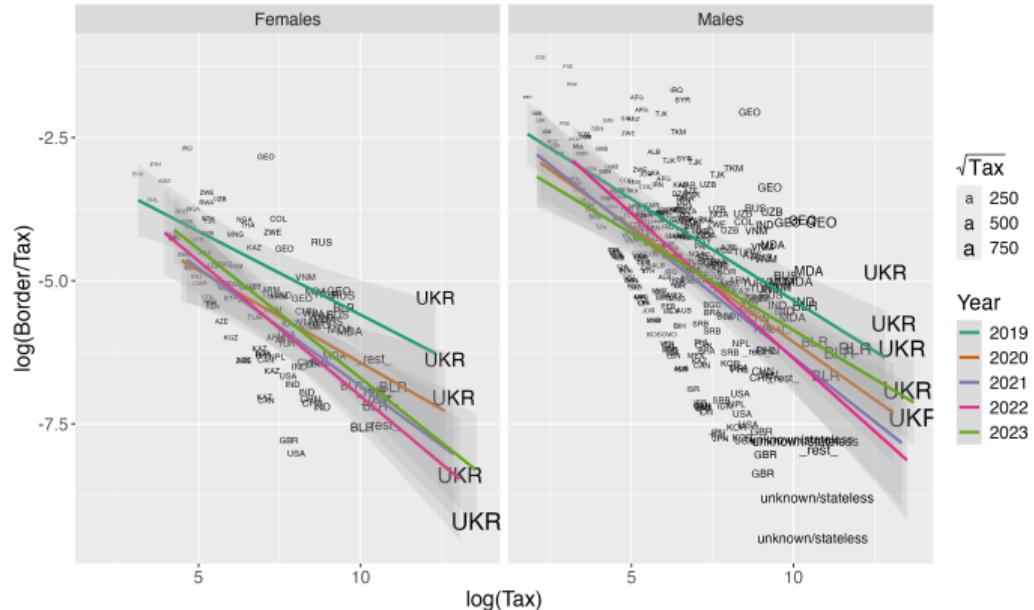


Figure 13: Relation between $\log(N)$ and $\log(m/N)$ between 2019 and 2023 and sexes

Assumptions – visual inspection (Tax register, by sex)

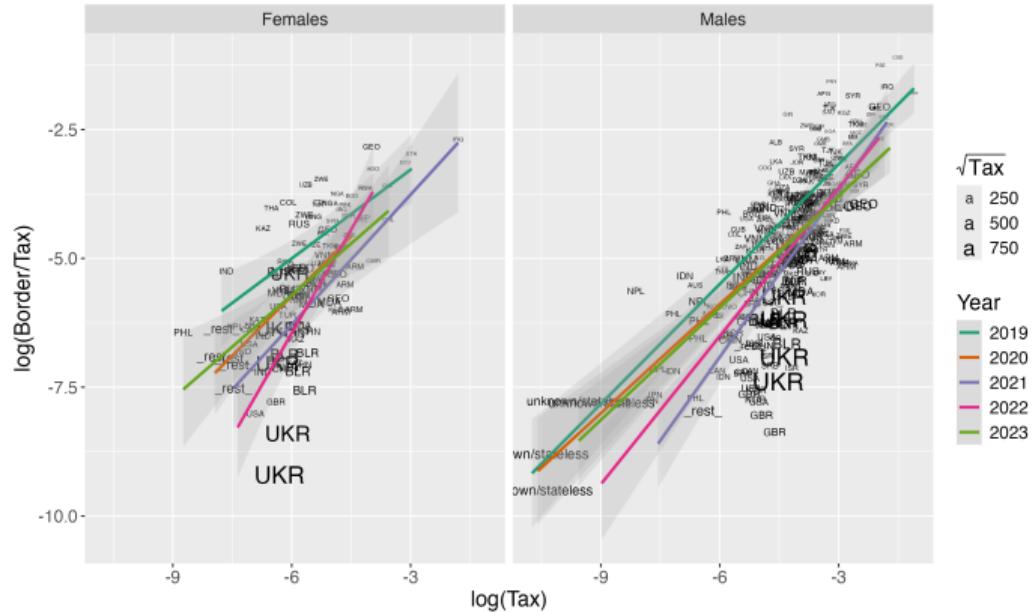


Figure 14: Relation between $\log(N)$ and $\log(m/N)$ between 2019 and 2023 and sexes

Assumptions – visual inspection (Population register, overall)

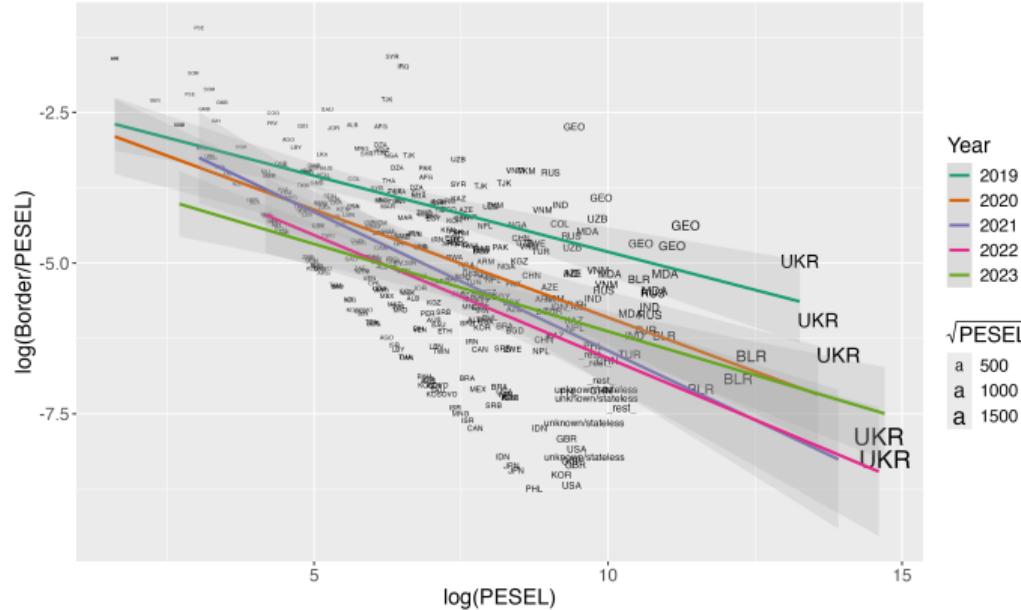


Figure 15: Relation between $\log(N)$ and $\log(m/N)$ between 2019 and 2023

Assumptions – visual inspection (Population register, overall)

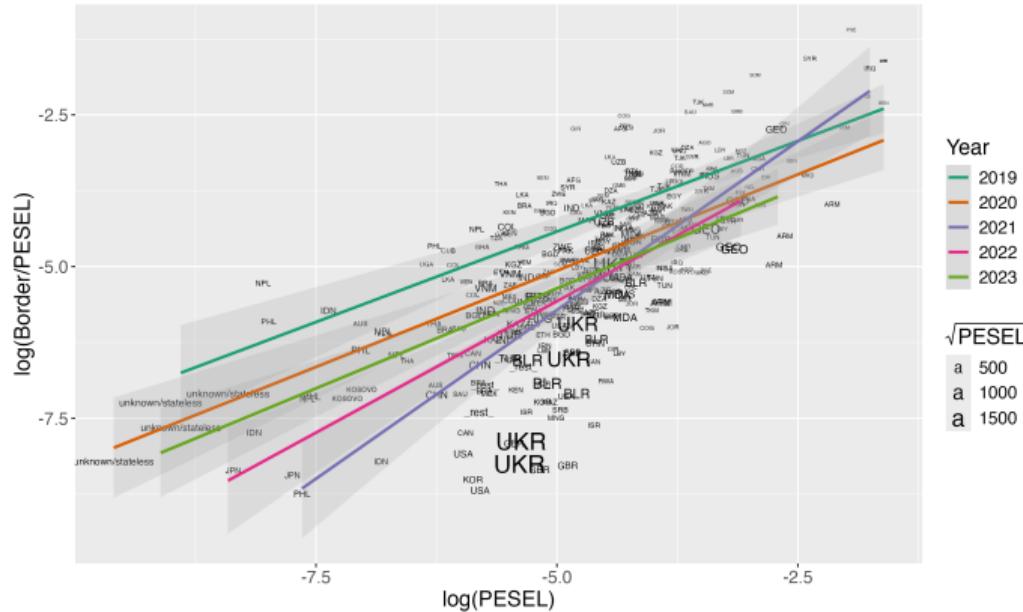


Figure 16: Relation between $\log(n/N)$ and $\log(m/N)$ between 2019 and 2023

Assumptions – visual inspection (Population register, by sex)

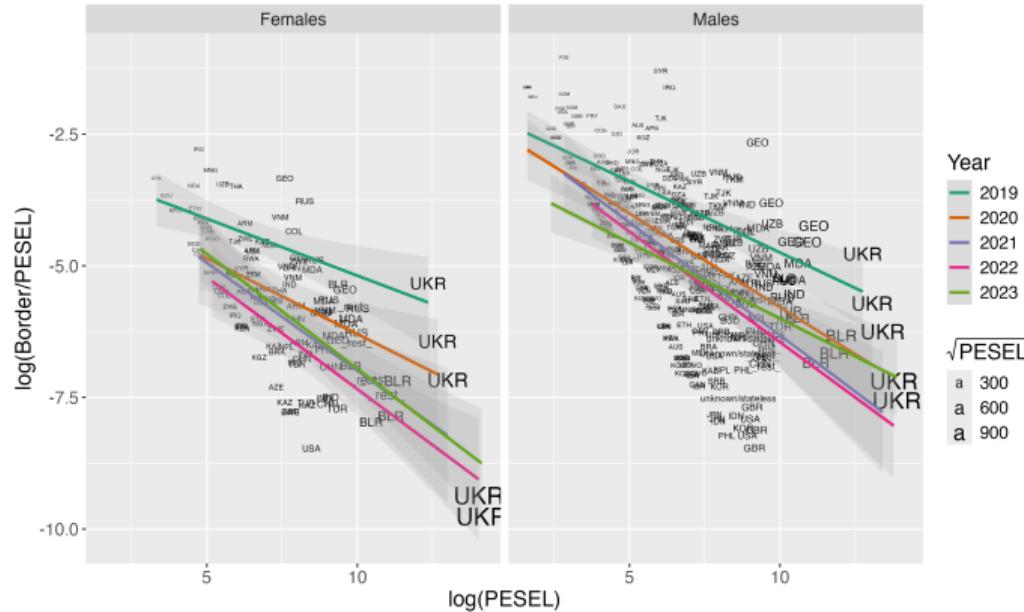


Figure 17: Relation between $\log(N)$ and $\log(m/N)$ between 2019 and 2023

Assumptions – visual inspection (Population register, by sex)

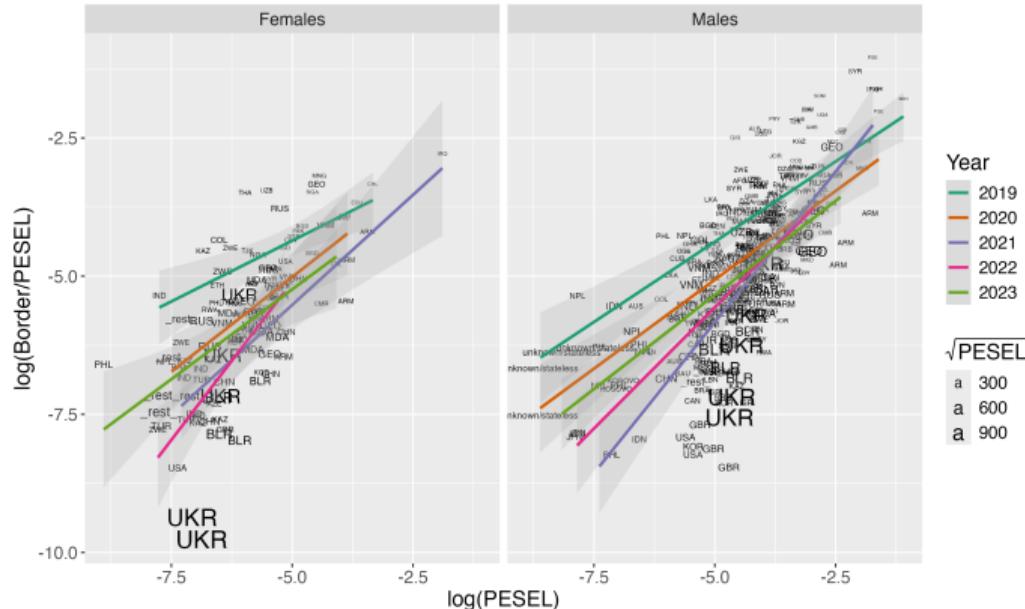


Figure 18: Relation between $\log(n/N)$ and $\log(m/N)$ between 2019 and 2023

Assumptions – visual inspection (Employees register, overall)

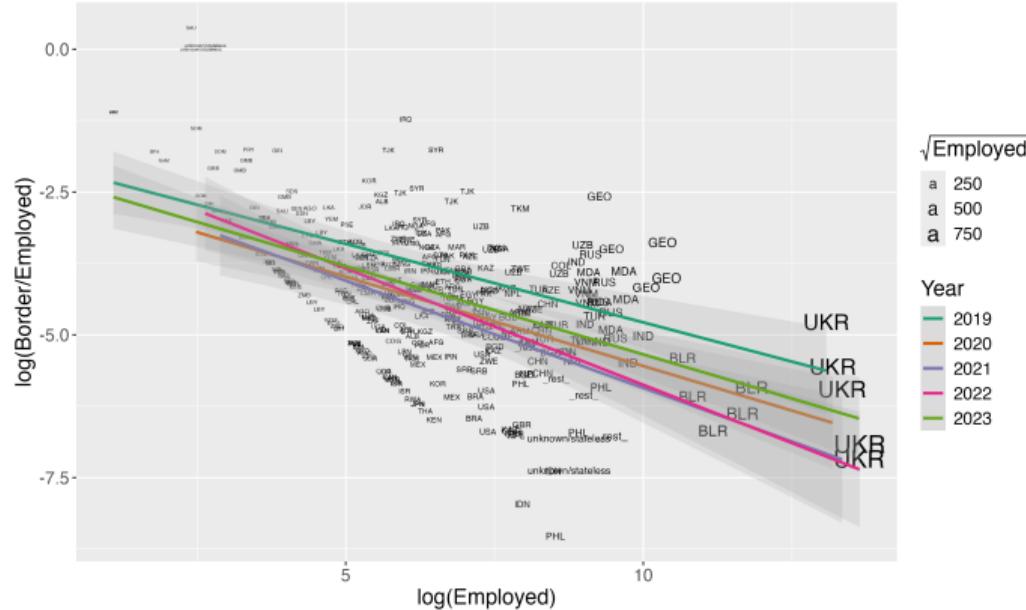


Figure 19: Relation between $\log(N)$ and $\log(m/N)$ between 2019 and 2023

Assumptions – visual inspection (Employees register, overall)

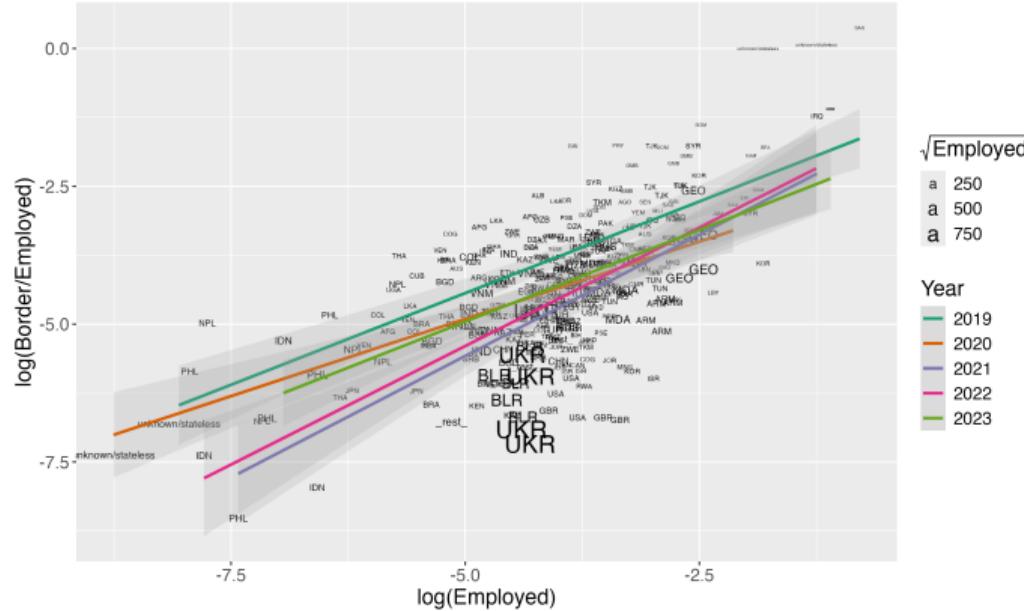


Figure 20: Relation between $\log(n/N)$ and $\log(m/N)$ between 2019 and 2023

Assumptions – visual inspection (Employees register, by sex)

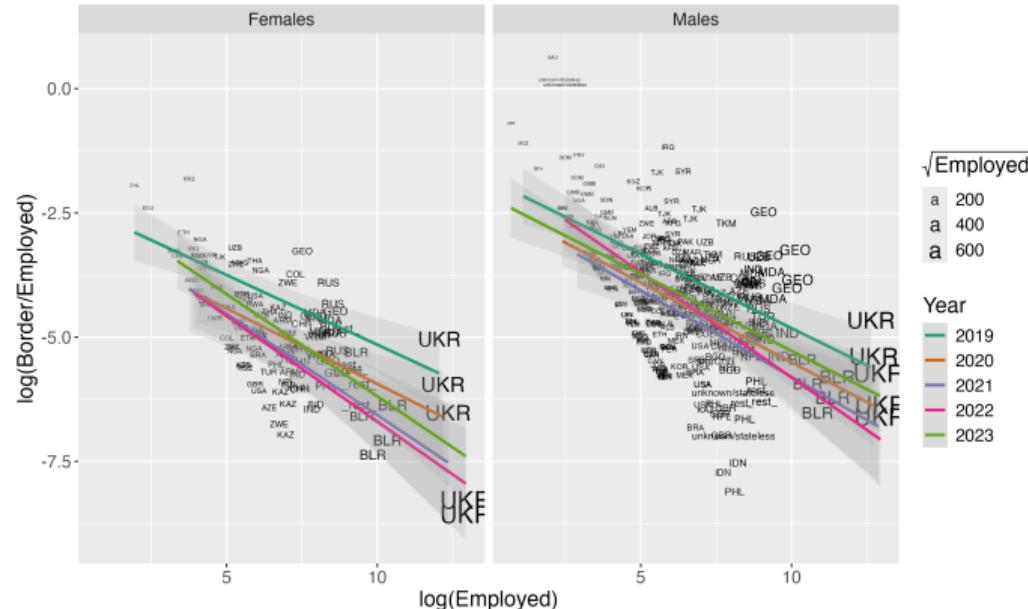


Figure 21: Relation between $\log(N)$ and $\log(m/N)$ between 2019 and 2023

Assumptions – visual inspection (Employees register, by sex)

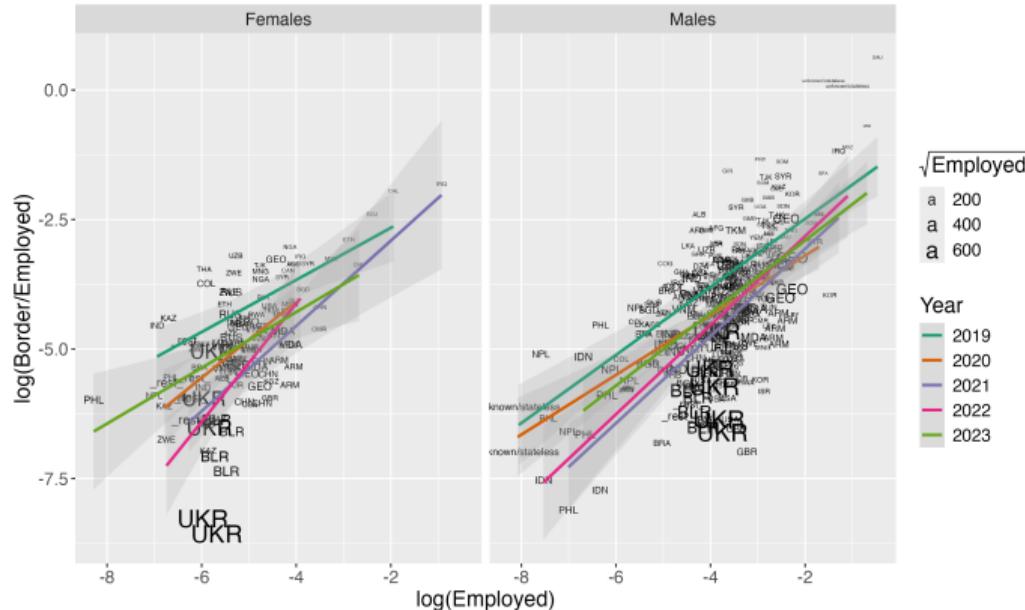


Figure 22: Relation between $\log(n/N)$ and $\log(m/N)$ between 2019 and 2023

8 Appendix

- Migration to Poland
- More about the data
- Assumptions – registers
- **Assumptions – registers by continent**
- Detailed results
- Sensitivity analysis

Assumptions – continents (Population register, by sex)

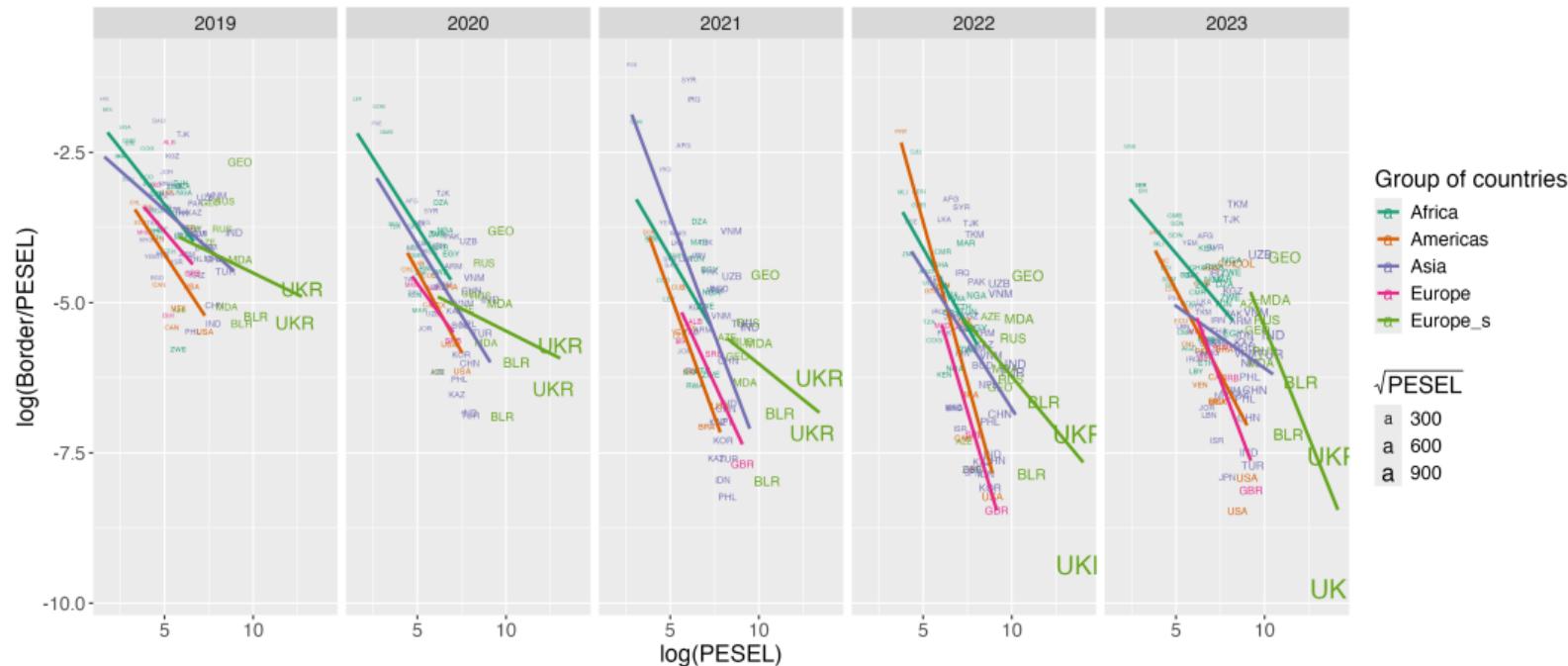


Figure 23: Relation between $\log(N)$ and $\log(m/N)$ between 2019 and 2023

Assumptions – continents (Population register, by sex)

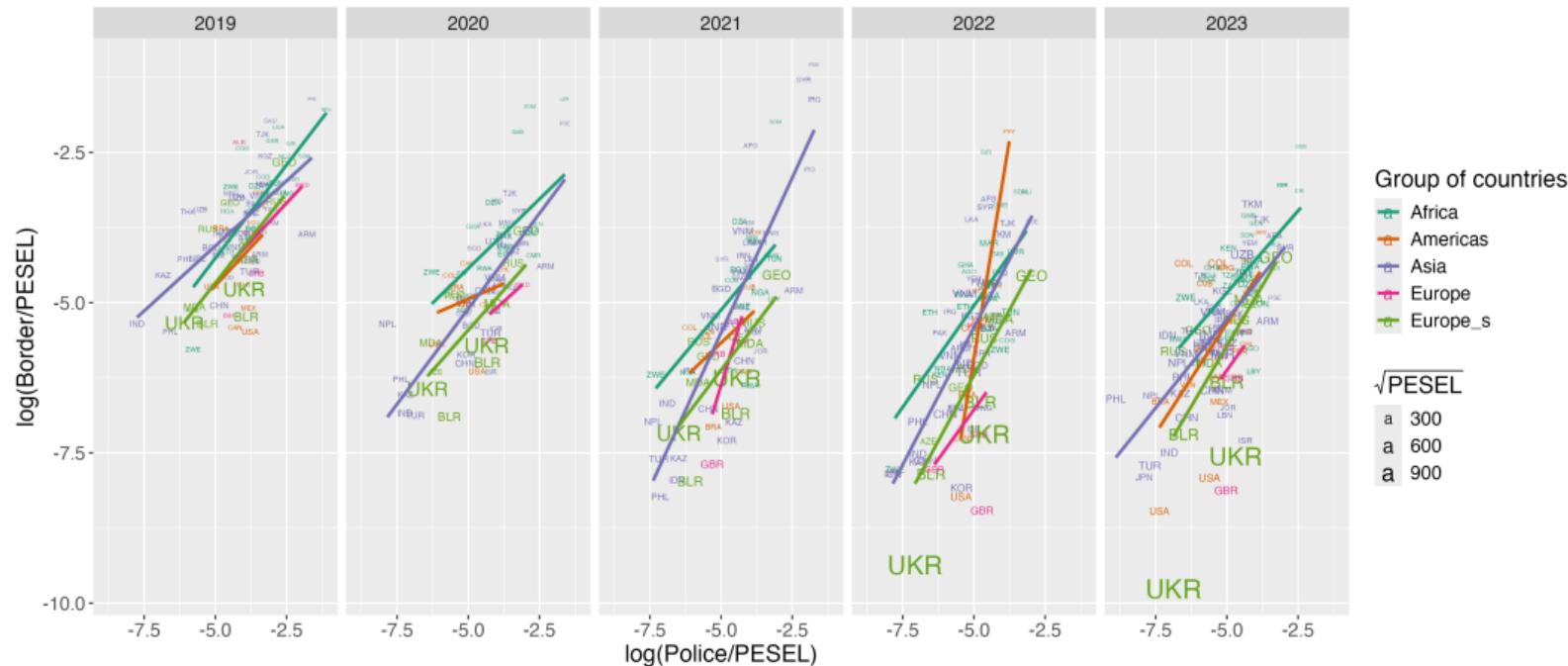


Figure 24: Relation between $\log(n/N)$ and $\log(m/N)$ between 2019 and 2023

Assumptions – continents (Tax register, by sex)

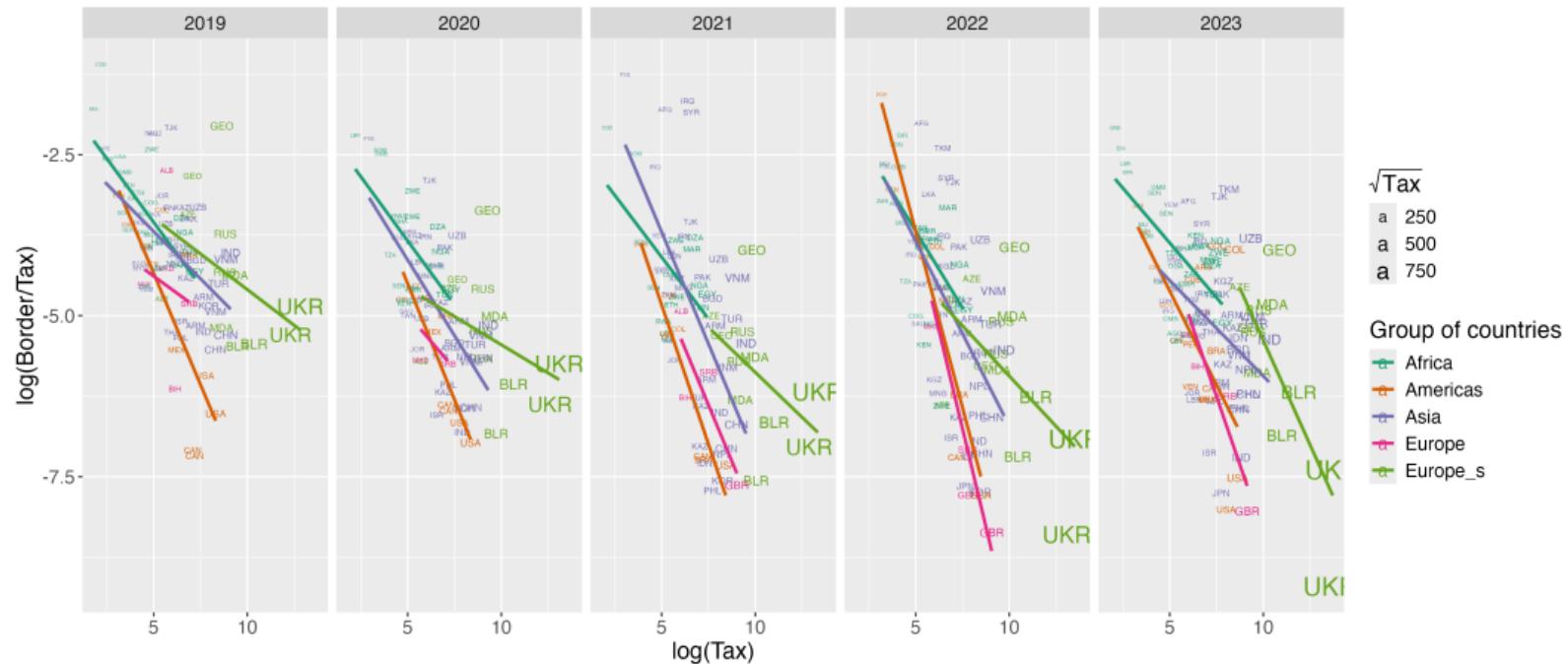


Figure 25: Relation between $\log(N)$ and $\log(m/N)$ between 2019 and 2023

Assumptions – continents (Tax register, by sex)

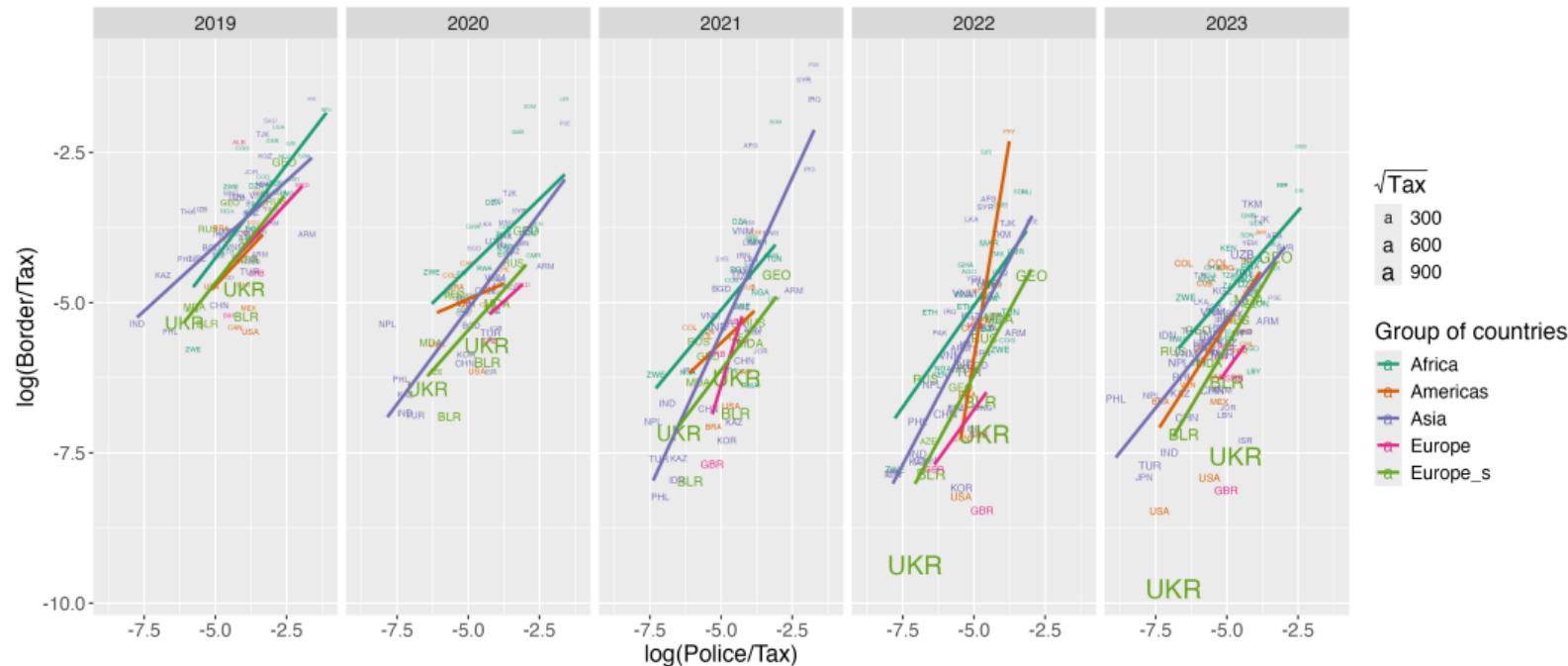


Figure 26: Relation between $\log(n/N)$ and $\log(m/N)$ between 2019 and 2023

Assumptions – continents (Employees register, by sex)

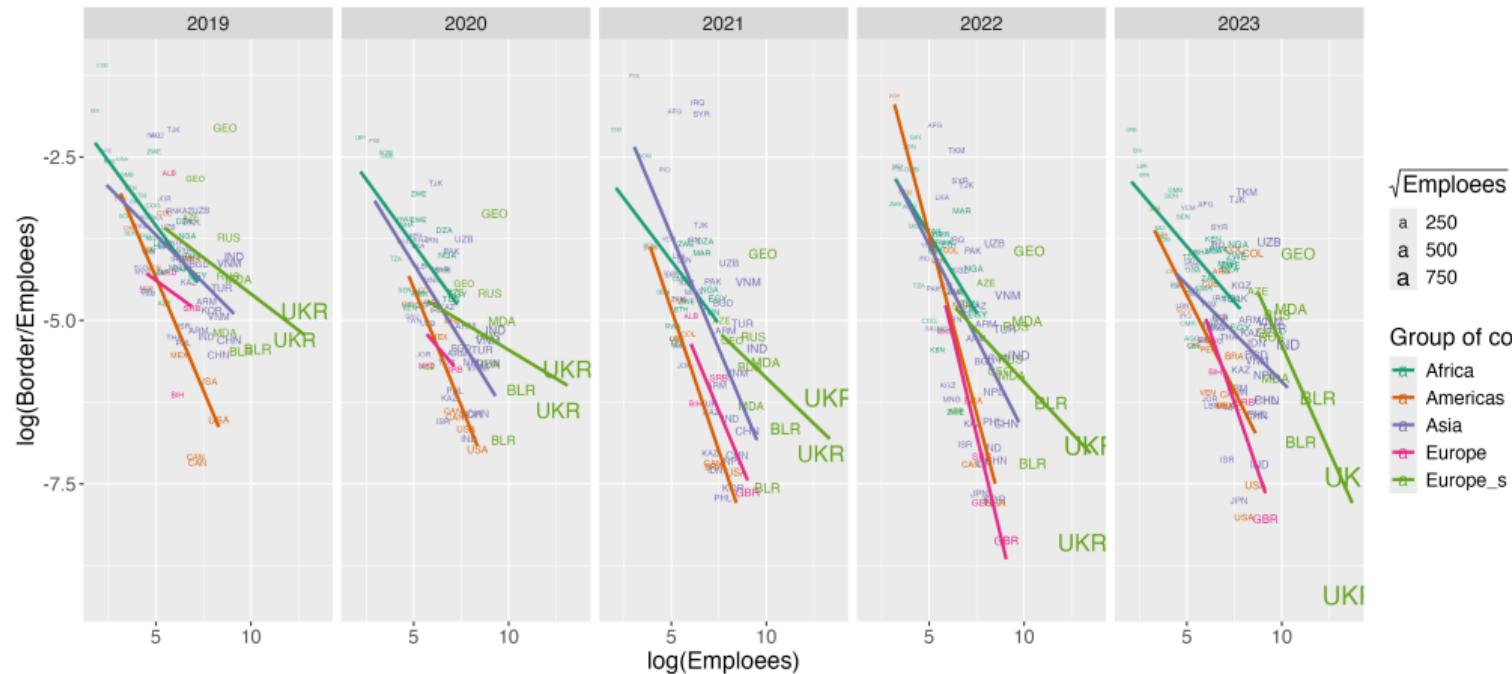


Figure 27: Relation between $\log(N)$ and $\log(m/N)$ between 2019 and 2023

Assumptions – continents (Employees register, by sex)

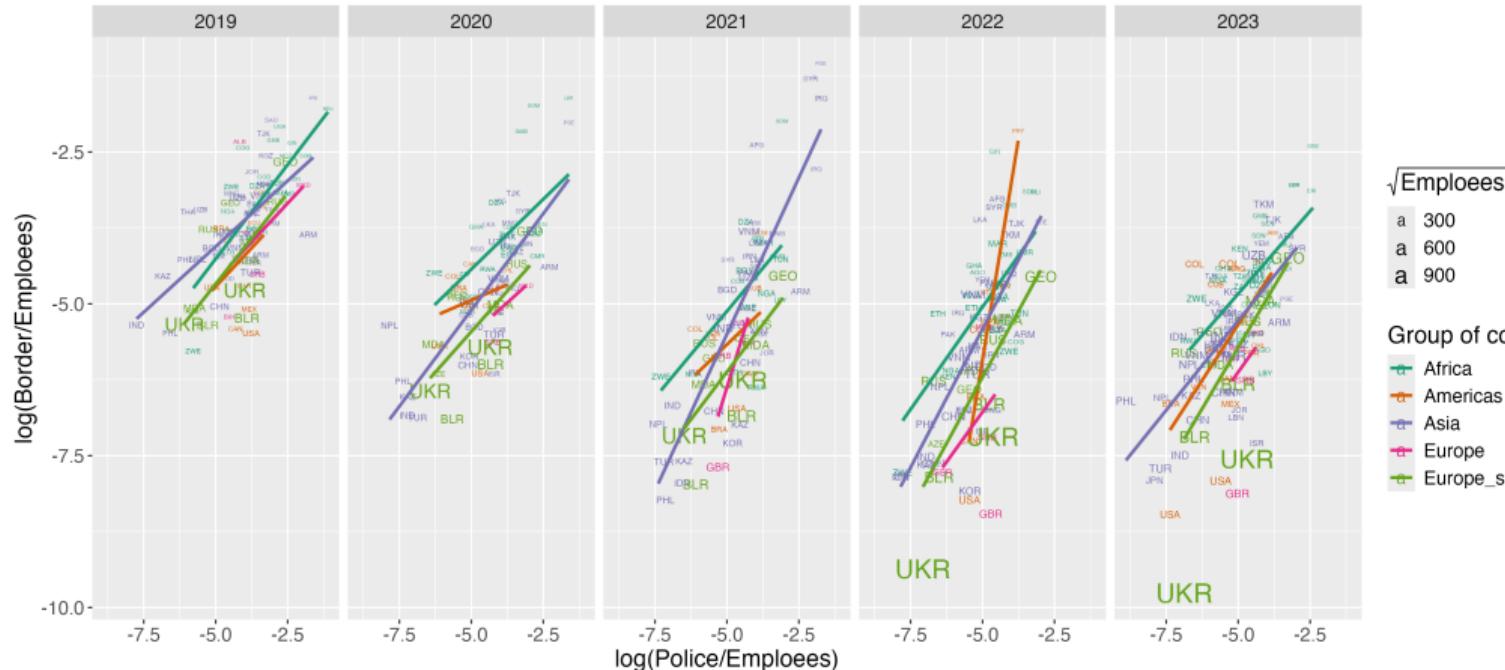


Figure 28: Relation between $\log(n/N)$ and $\log(m/N)$ between 2019 and 2023

8 Appendix

- Migration to Poland
- More about the data
- Assumptions – registers
- Assumptions – registers by continent
- **Detailed results**
- Sensitivity analysis

Estimates – single α (Tax register)

Table 6: Estimates of the size of the unauthorised population in Poland between 2019 and 2023 and its ratio to the reference population (tax register) based on three models with a single α

Year	OLS		Poisson		Negative Binomial	
	$\hat{\xi}$ (SE)	% of N (95% CI)	$\hat{\xi}$ (SE)	% of N (95% CI)	$\hat{\xi}$ (SE)	% of N (95% CI)
2019	37.7 (7.7)	4.5 (2.9, 6.4)	54.7 (25.2)	6.5 (3.1, 15.7)	52.7 (10.9)	6.2 (3.9, 9.1)
2020	43.4 (9.2)	4.1 (2.6, 5.9)	63.8 (30.5)	6.0 (2.8, 14.9)	61.3 (13.1)	5.7 (3.6, 8.5)
2021	51.0 (11.2)	3.7 (2.4, 5.4)	75.8 (37.5)	5.6 (2.6, 14.2)	72.7 (16.0)	5.3 (3.3, 8.0)
2022	59.8 (13.2)	3.6 (2.3, 5.3)	89.0 (44.4)	5.4 (2.5, 14.0)	85.4 (18.9)	5.2 (3.2, 7.8)
2023	80.3 (17.9)	3.4 (2.1, 5.0)	120.2 (61.4)	5.1 (2.3, 13.4)	115.3 (26.0)	4.9 (3.0, 7.4)

Note: $\hat{\xi}$ estimated using a model with a single α (and β) for the whole period. Bootstrap standard errors reported in brackets for the $\hat{\xi}$ estimates. For each approach, we report point estimates and 95% confidence intervals for the ratio to the reference population, calculated as $\hat{\xi}_t/N_t \times 100$ (same for the 95% CI).

Estimates – α varies over time (Tax register)

Table 7: Estimates of the size of the unauthorised population in Poland between 2019 and 2023 and its ratio to the reference population (tax register) based on three models with time-varying α

Year	OLS		Poisson		Negative Binomial	
	$\hat{\xi}$ (SE)	% of N (95% CI)	$\hat{\xi}$ (SE)	% of N (95% CI)	$\hat{\xi}$ (SE)	% of N (95% CI)
2019	91.9 (20.7)	10.9 (6.8, 16.0)	116.3 (32.1)	13.8 (8.8, 22.2)	114.9 (25.1)	13.6 (9.0, 20.8)
2020	51.4 (11.9)	4.8 (3.0, 7.6)	71.7 (17.0)	6.7 (4.2, 10.2)	58.4 (12.4)	5.5 (3.6, 8.3)
2021	35.4 (8.7)	2.6 (1.6, 4.4)	59.1 (14.5)	4.3 (2.6, 6.4)	51.1 (12.5)	3.7 (2.4, 6.2)
2022	48.5 (11.1)	3.0 (1.9, 4.5)	50.6 (15.0)	3.1 (1.9, 4.9)	65.4 (14.1)	4.0 (2.7, 6.1)
2023	99.7 (23.6)	4.3 (2.6, 6.7)	85.6 (41.0)	3.7 (1.8, 7.6)	137.4 (32.3)	5.9 (3.8, 9.5)

Note: $\hat{\xi}$ estimated using a model with time-varying α (and a single β for the whole period). Bootstrap standard errors reported in brackets for the $\hat{\xi}$ estimates. For each approach, we report point estimates and 95% confidence intervals as percentages of the reference population, calculated as $\hat{\xi}_t/N_t \times 100$.

Estimates – α varies over time and between sex (Tax register)

Table 8: Estimates of the size of the unauthorised population in Poland between 2019 and 2023 and its ratio to the reference population (tax register) based on three models with time-varying and sex-specific α

Year	OLS		Poisson		Negative Binomial	
	$\hat{\xi}$ (SE)	% of N (95% CI)	$\hat{\xi}$ (SE)	% of N (95% CI)	$\hat{\xi}$ (SE)	% of N (95% CI)
2019	121.0 (30.5)	14.3 (8.5, 21.6)	172.4 (54.2)	20.4 (12.0, 38.3)	136.2 (33.1)	16.1 (9.7, 24.6)
2020	66.9 (17.2)	6.3 (3.8, 9.8)	103.3 (32.6)	9.7 (5.2, 17.3)	68.1 (13.9)	6.4 (4.1, 9.3)
2021	44.3 (11.1)	3.2 (1.9, 4.9)	79.0 (23.1)	5.8 (3.4, 10.4)	56.6 (13.6)	4.1 (2.7, 6.5)
2022	58.1 (14.7)	3.5 (2.0, 5.4)	59.5 (20.6)	3.6 (2.2, 5.9)	71.5 (15.6)	4.4 (2.8, 6.5)
2023	128.7 (38.3)	5.5 (2.7, 8.8)	100.6 (65.8)	4.3 (1.8, 9.2)	163.9 (44.4)	7.0 (4.2, 11.3)

Note: $\hat{\xi}$ estimated using a model with α varying over time and sex (and a single β for the whole period). Bootstrap standard errors reported in brackets for the $\hat{\xi}$ estimates. For each approach we report point and 95% confidence relation to the reference population calculated as $\hat{\xi}_t/N_t \times 100$ (same for 95% CI).

8 Appendix

- Migration to Poland
- More about the data
- Assumptions – registers
- Assumptions – registers by continent
- Detailed results
- Sensitivity analysis

Alternative covariate: prison

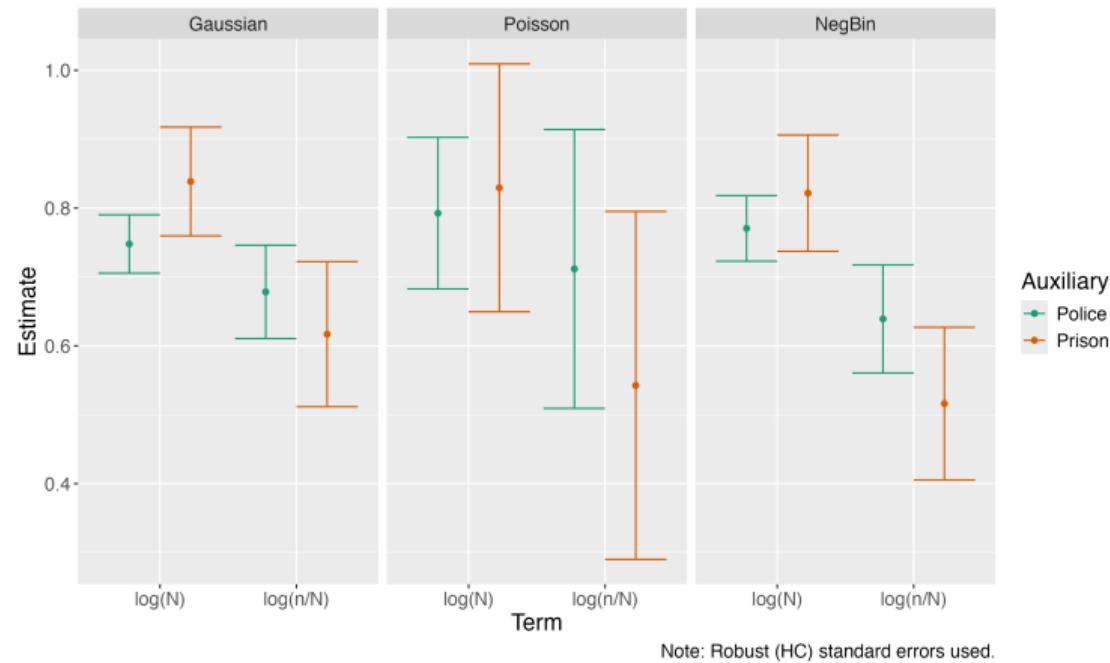


Figure 29: Sensitivity analysis for the α and β for OLS, Poisson and Negative Binomial

Alternative covariate: prison and time varying α 's

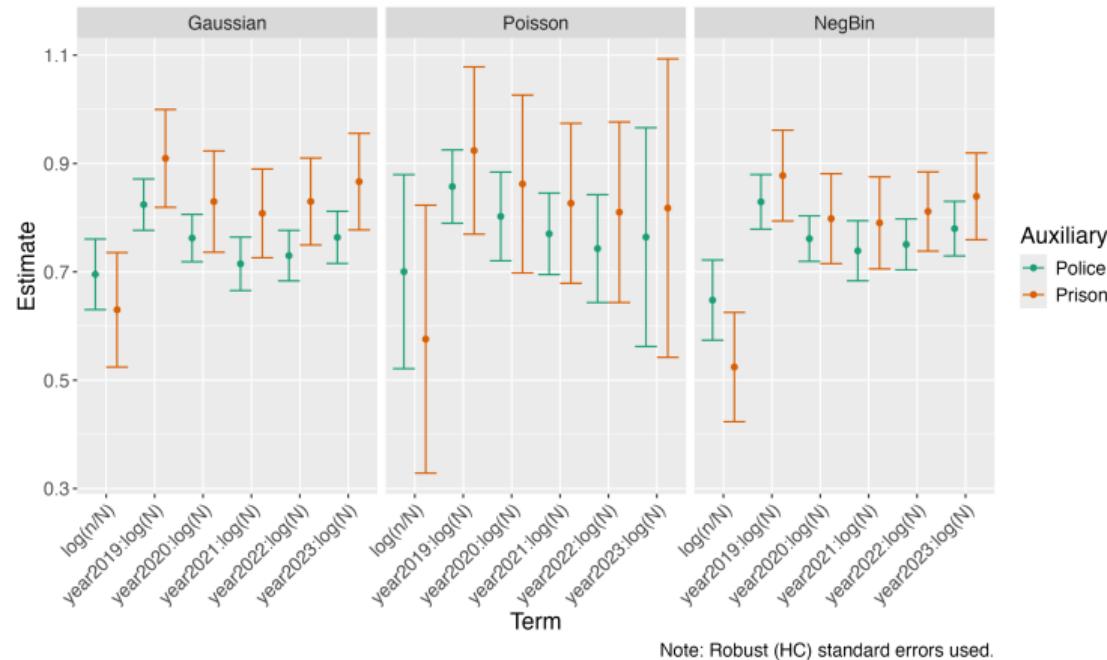


Figure 30: Sensitivity analysis for the time varying α 's and a single β for OLS, Poisson and Negative Binomial