Assessing the spatial overlap between urban fragmentation and residential segregation in European cities.

Lucas Spierenburg[1*], Esteban Ralon-Santizo[2], Sander van Cranenburgh[2], Oded Cats[1]

[1]Transport and Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, Delft, 2628CN, the Netherlands.

[2]Transport and Logistics Group, Department of Engineering Systems and Services, Faculty of Technology, Policy, and Management, Delft University of Technology, Jaffalaan 5 , Delft, 2628BX, the Netherlands.

*Corresponding author(s). E-mail(s): l.j.spierenburg@tudelft.nl; Contributing authors: e.d.ralonsantizo@student.tudelft.nl; s.vancranenburgh@tudelft.nl; o.cats@tudelft.nl;

1 Main

Residential segregation, defined as the uneven distribution of social groups in space, has been consistently linked to disparate outcomes in education, employment, and health for disadvantaged populations [1–5]. Understanding the mechanisms that produce and perpetuate this phenomenon is therefore crucial for addressing its societal consequences. Research has long documented how residential segregation patterns emerge alongside urban development [6]. In particular, scholars have found that residential segregation often overlaps with urban fragmentation — i.e., the partition of urban space through urban barriers such as railways, motorways, and waterways — where urban barriers act as frontiers between social groups [5, 7].

The phenomenon that residential segregation aligns with urban fragmentation has been extensively demonstrated in the United States and South Africa, where urban barriers clearly separate social groups—epitomised by the common expression "wrong side of the tracks" [6, 8–12]. In these countries, alignment between residential segregation and urban fragmentation was policy-driven, with urban barriers serving as tools for enforcing social separation through explicit segregation policies like redlining in the US and apartheid in South Africa [13–16]. Theory suggests that alignment between residential segregation and urban fragmentation could also occur naturally from people's residential choice patterns without explicit separation policies due to homophily, i.e. the preference to be around people similar to oneself [17, 18]. In this context, urban barriers would allow groups to separate from dissimilar others.

While alignment between residential segregation and urban fragmentation in contexts with explicit segregation policies has been observed, and theory provides arguments that such alignment could also emerge naturally, empirical evidence is lacking regarding it happening in contexts without segregation-encouraging policies. Addressing this knowledge gap would yield significant theoretical and practical implications. Theoretically, it would test whether such alignment can emerge solely from residential preferences without policy-driven segregation. From a policy perspective, it would clarify whether removing urban barriers that fragment space would help in reducing residential segregation.

We address this knowledge gap by quantifying how residential segregation patterns overlap with the fragmentation of the urban space across Europe. Our study encompasses 520 cities with populations exceeding 50,000 across eight Western European countries: Germany, France, Ireland, Italy, the Netherlands, Portugal, Spain, and the United Kingdom. We use the high-resolution gridded dataset of migrant populations developed by the EU's Joint Research Centre [19]. We focus specifically on the residential segregation of non-EU migrants and analyze whether residential segregation aligns with urban fragmentation more than would be expected by chance.[1] Our methodological framework employs a Monte Carlo approach, in which we generate 200 synthetic fragmentation patterns for each city and compare the actual spatial overlap against this null distribution. Figure 1 summarizes our analysis approach. We identify demographically homogeneous regions using a regionalization method that aggregates grid cells based on demographic data [20]. Urban fragments are defined as contiguous groups of cells bounded by railways, motorways, or waterways. We test the statistical significance of the overlap between the demographic partition and the fragmentation partition by comparing the observed value against 200 synthetic fragmentation partitions.

Our results challenge the assumption that urban barriers function as social frontiers in the absence of segregation-encouraging policies. Specifically, we do not observe a systematic alignment between residential segregation patterns and urban fragmentation across European cities. In fact, we detect a statistically significant pattern in the opposite direction, with segregation patterns in many cities showing less alignment than would have been expected by chance. Interestingly, we identify important regional variations, with cities in Germany and the Netherlands exhibiting significantly stronger alignment between barriers and segregation compared to other countries in our sample, even after controlling for city-specific characteristics. These findings suggest that

[1] The United Kingdom was part of the European Union in 2011, the time at which the data was produced.

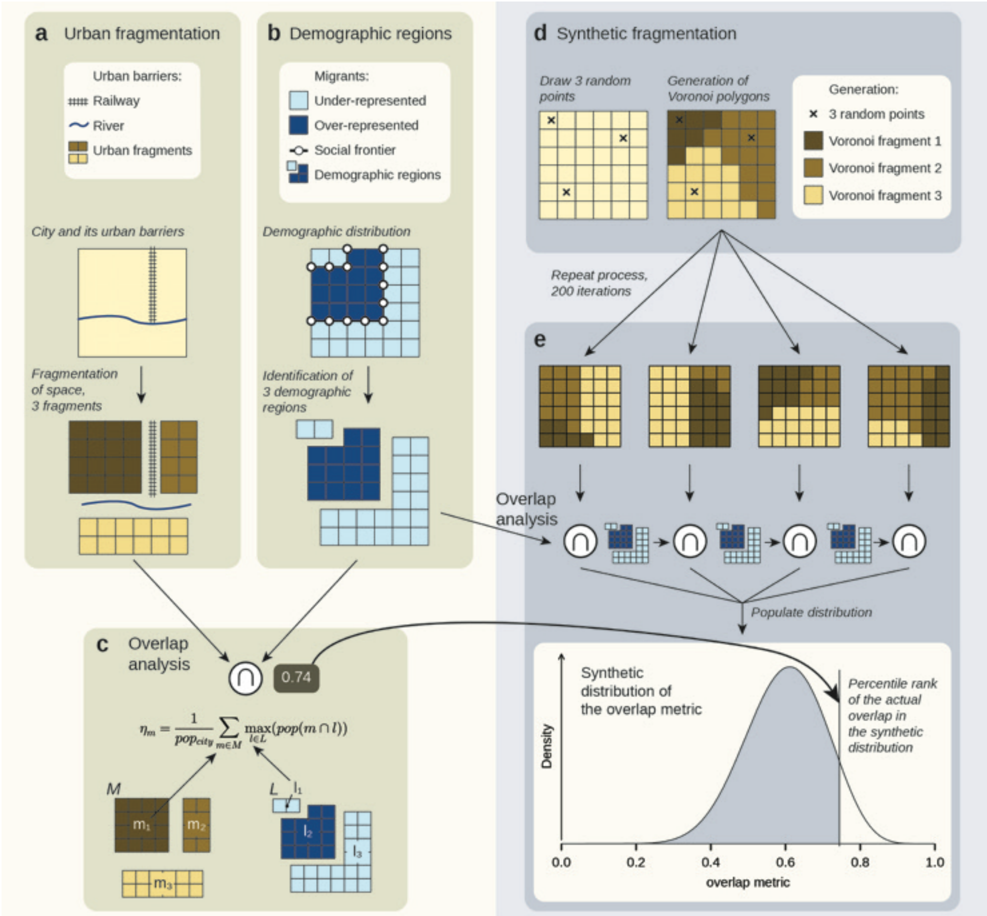

Fig. 1 Quantifying the statistical significance of the overlap between residential segregation and urban fragmentation at the city level. (a) Fragmentation of urban space by urban barriers into three fragments. (b) Spatial partitioning of space into demographically homogeneous regions. (c) Quantifying the overlap between residential segregation and urban fragmentation, using the purity score. (d) Generating 200 synthetic urban fragmentation partitions, using Voronoi polygons. (e) Computing the overlap between residential segregation and each synthetic fragmentation partition, generating the statistical distribution and evaluating the quantile rank of the actual overlap in the distribution of the synthetic overlaps.

the relationship between urban barriers and social frontiers is neither universal nor random, but rather influenced by country-specific urban development patterns.

2 Results

Our analysis of 520 cities across eight European countries reveals two key findings regarding the spatial relationship between residential segregation patterns and urban fragmentation. First, contrary to expectations based on findings reported for

North American cities, we do not observe a systematic alignment between residential segregation and urban fragmentation across European cities (subsection 2.1). Second, the cities that do exhibit significant alignment are predominantly concentrated in the Netherlands and Germany, suggesting important regional variations in this relationship (subsection 2.2).

2.1 No alignment between residential segregation and urban fragmentation in European cities

To assess the extent to which residential segregation aligns with urban fragmentation in the European context, we investigate the statistical distribution of quantile ranks for all 520 urban areas in our sample (Figure 2 a). Here, urban barriers refer to railways, motorways, and waterways that fragment urban space, while social frontiers denote the boundaries between demographically distinct regions. If residential segregation aligns with urban fragmentation across European cities, we would expect an over-representation of cities in the right portion of the histogram, particularly in the highest quantile rank bin (0.95-1.00), indicating that urban barriers align with social frontiers to a degree unlikely to occur by chance. Conversely, if no association exists, we would anticipate a uniform distribution of quantile ranks across all bins. The observed distribution reveals a pattern that is predominantly uniform with a slight L-shape, characterized by higher counts in the first (0.00-0.05) bin compared to the rest (Figure 2 a). Importantly, in the vast majority of the cities that we investigate (N=486, 93.5%), residential segregation patterns show no significant alignment to urban fragmentation, with their quantile rank falling outside the top 5% of the synthetic distribution.

To assess whether the observed distribution pattern deviates significantly from what would be expected under random chance, we use a formal statistical testing framework that accounts for multiple comparisons using Benjamini-Yekutieli adjusted p-values (Figure 2 a) [21]. Our statistical test reveals that only the lowest quantile rank bin (0.00-0.05) shows a statistically significant deviation from expected counts, with cities being overrepresented in this category. This finding suggests a pattern that is opposite to the expected alignment between urban barriers and social frontiers, with a small subset of European cities exhibiting less alignment between urban barriers and social frontiers than what would be expected by chance. This result is contrary to the expectation that urban barriers delineate segregation patterns across European urban contexts, highlighting a fundamental difference in patterns observed in their North American counterparts.

To shed more light on this result, we showcase the results from two cities that exhibit clear disagreement between residential segregation and urban fragmentation: Lyon (France) and Birmingham (United Kingdom), shown in figure 3. In both cities, we observe distinct spatial patterns in the demographic distribution. While Birmingham exhibits a center-periphery pattern with migrant populations being overrepresented in the center (figure 3 a), Lyon displays a pronounced East-West division (figure 3 b), with migrants predominantly residing in eastern districts. Despite the presence

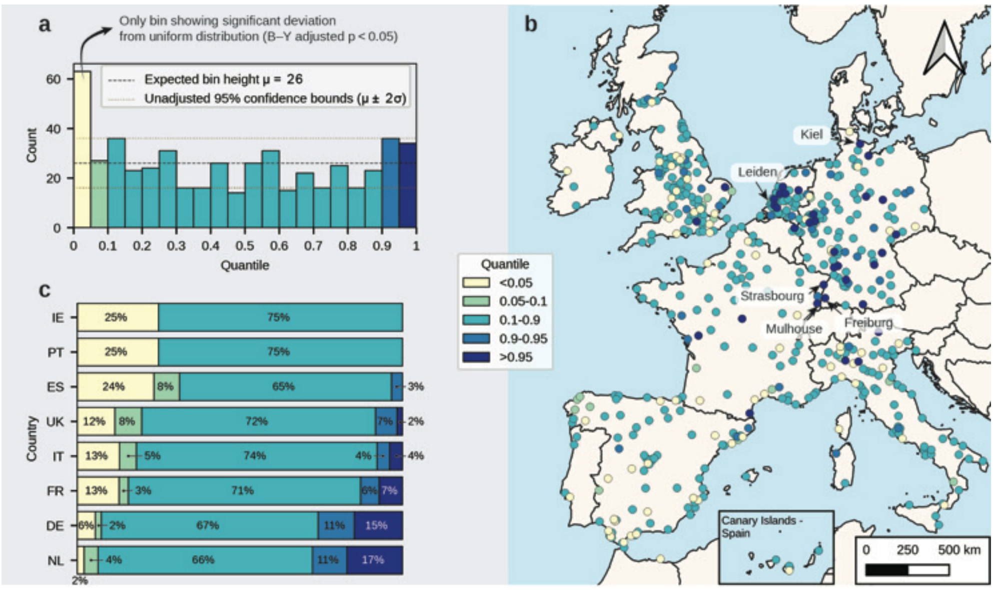

Fig. 2 Statistical and geographical patterns of alignment between residential segregation and urban fragmentation across European cities. (a) Distribution of quantile ranks across all cities in the dataset. Each bar shows the number of cities falling within a specific quantile rank bin, while the dashes horizontal line indicates the expected count per bin under a null scenario where the alignment between residential segregation patterns and urban fragmentation is not substantially different from what would be expected by chance (no systematic relationship exists between residential segregation and urban fragmentation). The dotted horizontal lines indicate significance thresholds without applying the Benjamini-Yekutieli correction. This correction accounts for the inherent dependence structure of the data. (b) Geographic distribution of quantile ranks across cities in the dataset. Cities mentionned in the paper are (c) Distribution of quantile ranks across countries.

of clear social frontiers, our analysis reveals minimal alignment with urban barriers in both contexts. In Lyon, the *boulevard périphérique* (D383) follows a similar East-West orientation as the social frontier; nevertheless, rather than marking the frontier between demographic regions, it runs through the core of the area over-representing migrants, functioning more as a "spine" for this region rather than as a dividing barrier. Similarly, Birmingham's center contains dense infrastructure forming a central node in a radial network, yet these physical elements do not delineate social frontiers between areas over-representing migrants and the rest of the city. These two cities are illustrations of the statistical pattern observed across our sample, demonstrating that even in cities with pronounced segregation patterns, urban barriers do not necessarily systematically define social frontiers in the European context.

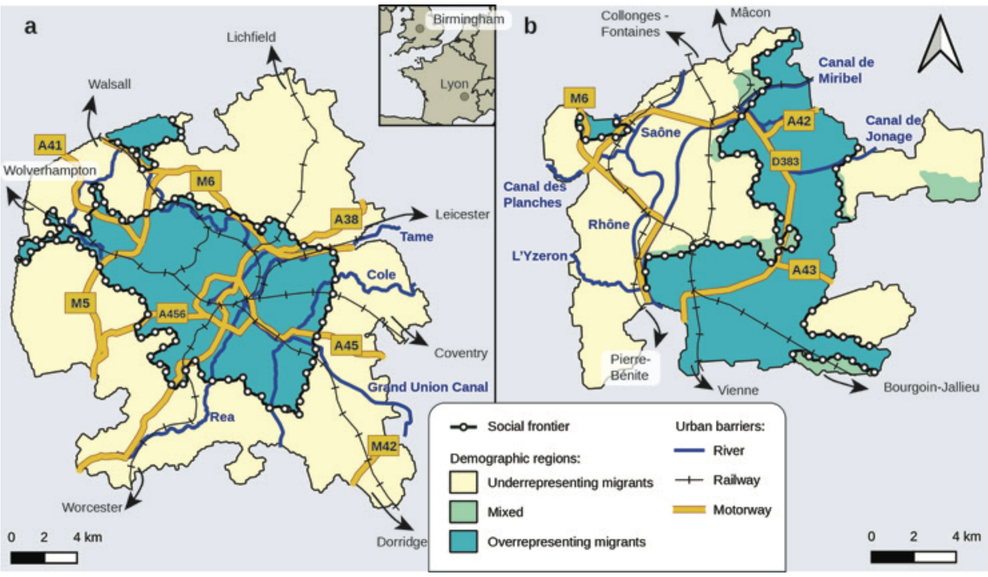

Fig. 3 Spatial layout of social frontiers and urban barriers in Birmingham, United Kingdom (a) and in Lyon, France (b). Selected highways, rivers and railways are displayed to enhance readability.

2.2 Greater association between residential segregation and urban fragmentation in Germany and the Netherlands than in other countries

We identify geographical differences in the extent to which social frontiers align with urban barriers throughout Europe (Figure 2). While we do not observe a consistent alignment between residential segregation patterns and the urban fragmentation phenomenon in our sample, urban areas exhibiting such alignment are predominantly concentrated in Germany and the Netherlands. There exists stronger alignment between residential segregation patterns and urban fragmentation in a subset of cities in Germany (15 % of cities) and the Netherlands (17 % of cities). This pattern suggests that the relationship between urban barriers and social frontiers may be influenced by country-specific urban development strategies. Yet, even in Germany and the Netherlands, cases exhibiting significant correspondence between urban barriers and demographic frontiers constitute a minority of cities in these countries. Interestingly, the French region of Alsace, where cities like Strasbourg and Mulhouse were administered by Germany between 1870 and 1918, also shows significant alignment between urban fragments and demographic regions.

To formally assess the potential significance of these country-specific effects, we conduct a beta regression analysis controlling for two city-level variables that might explain these geographical differences [22]. We control for segregation intensity using the dissimilarity index (subsection 4.3.2), as the alignment between segregation patterns and urban fragmentation might be less observable in cities where segregation is less pronounced. We also control for decentralization, as multicentric urban structures may be more likely to exhibit barrier-frontier alignment because physical barriers can serve as natural separators between distinct development centers, whereas monocentric cities with radial growth patterns tend to extend across barriers rather than being bounded by them. We use a normalized average distance from city center for measuring decentralization (see subsection 4.3.2). Figure 4 displays the distribution of segregation intensity and decentralization across the countries in our sample. Spain and Italy exhibit the highest levels of segregation intensity, while Portugal and France show the lowest, consistent with findings from [23] using the same data and dissimilarity indicator. Germany and the Netherlands exhibit moderate levels of segregation intensity relative to other countries in the sample. The decentralization indicator, on the contrary, does show a similar geographic pattern to that of the alignment between physical barriers and social frontiers, with Germany and the Netherlands showing the highest levels of spatial decentralization, while Spain and Portugal display the most centralized urban structures.

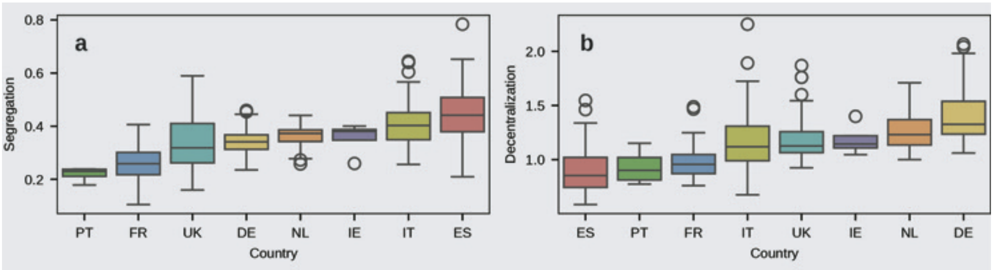

Fig. 4 Cross-country variation in residential segregation intensity and urban spatial structure. (a) Distribution of residential segregation across urban areas per country, measured using the dissimilarity index. (b) Distribution of the decentralization indicator across urban areas per country. The countries are arranged from left to right by ascending median.

The regression model estimation results show that Dutch cities have a significant positive association (average quantile rank of 0.61 p=0.008), even after controlling for decentralization, while Germany's association approaches but does not reach the 5% significance threshold (average quantile rank of 0.56 p=0.064). In contrast, we observe significant negative association in Spain (average quantile rank of 0.34 p<0.001), the United Kingdom (average quantile rank of 0.41 p=0.001), and Italy (average quantile rank of 0.43 p=0.03), reinforcing our earlier findings. These geographic disparities suggest systematic country-level differences besides decentralization in how residential segregation patterns relate to urban barriers.

Examining our control variables, we find that segregation intensity is not statistically significant, indicating that we cannot attribute the lack of observed alignment to an insufficient level of segregation. Conversely, decentralization shows a significant positive effect (p=0.023), suggesting that the spatial organization of cities plays an

	coef (logit scale)	p-value	Predicted average
Segregation	-0.1042	0.869	-
Decentralization	0.5759	0.022	-
Portugal	-0.9711	0.093	0.27
Spain	-0.6844	0.000	0.34
Ireland	-0.5049	0.399	0.38
the United Kingdom	-0.3605	0.001	0.41
Italy	-0.2979	0.03	0.43
France	-0.0645	0.687	0.48
Germany	0.2393	0.064	0.56
the Netherlands	0.4567	0.008	0.61

Table 1 Regression model estimation results. The model includes country-specific fixed effects, allowing us to identify countries where the alignment is significantly different from the null distribution. A positive coefficient implies a positive impact of the variable on the quantile rank. The variables for segregation and decentralization are centered before the regression, subtracting their respective mean. The average predicted quantile rank per country is measured using the transformation $logit^{-1}(\alpha_c)$ and provided in the last column of the table.

important role in the alignment between urban barriers and social frontiers. This finding aligns with the distinction between monocentric cities prevalent in countries like Spain and France versus the more multicentric urban agglomeration structures characteristic of Germany and the Netherlands (figure 4). The latter configuration may facilitate a more systematic division of space compared to the predominantly radial development patterns of monocentric cities.

To illustrate our statistical findings with concrete examples, we conduct a qualitative analysis of Amsterdam (Figure 5). Amsterdam is one of the cities exhibiting the highest quantile rank values (quantile rank =1) in our dataset, indicating exceptional alignment between residential segregation patterns and urban fragmentation. In Amsterdam, five main regions have an over-representation of migrants: Nieuw-West, Noord, Oost, along the Beneluxbaan, and Zuidoost. Crucially, these regions are delineated with remarkable correspondence to the city's waterways. In the western part of the city, a succession of canals (from Westlandgracht to Westelijk Marktkanaal) clearly separates a region over-representing migrants in the West from an area under-representing migrants in the East of the canals (figure 5 b). Similarly, in Amsterdam Noord (figure 5 c), the major canal of Ringsloot creates an almost perfect boundary between the northern area over-representing migrants and the southern area under-representing migrants. Similar striking overlap (quantile rank > 0.95) occurs in several Dutch and German cities, including Leiden (quantile rank $= 0.97$), Freiburg (quantile rank $= 0.97$), and Kiel (quantile rank > 0.995) (figure 2).

3 Discussion

Our analysis of 520 European cities challenges the assumption that residential segregation patterns align with urban fragmentation. This finding represents a significant departure from previous research, which has largely focused on contexts where explicit segregation policies were historically implemented, such as cities affected by redlining

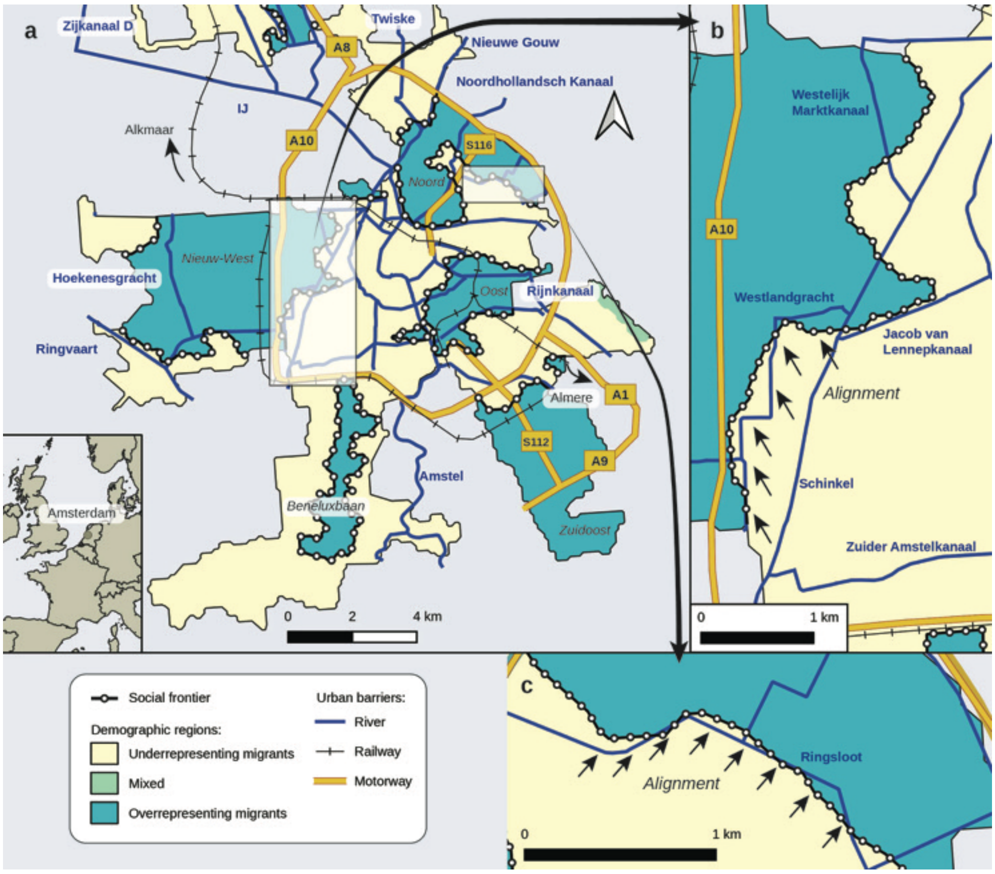

Fig. 5 Spatial layout of social frontiers and urban barriers in Amsterdam, the Netherlands. (a) Entire map of Amsterdam. (b) Zoom over Amsterdam West. (c) Zoom over Ringsloot in Amsterdam Noord. Selected highways, rivers and railways are displayed to enhance readability.

in the United States and the apartheid in South Africa [8, 15, 16, 24, 25]. Our findings demonstrate that the same pattern does not emerge when such explicit policies are absent. While households may indeed seek separation and sort themselves into homogeneous neighborhoods, this sorting does not appear to systematically occur along urban barriers in the European urban context.

In fact, we observe the opposite in a significant number of cities where residential segregation aligns less with urban fragmentation than expected under the null hypothesis. This suggests that when groups sort residentially near urban barriers, other factors may outweigh homophilic preferences for separation from dissimilar groups. Urban barriers may generate effects that either attract or repel certain population groups. For instance, areas adjacent to motorways experience higher noise and air pollution on both sides, making these locations less desirable regardless of which side of the barrier they occupy, thus potentially concentrating disadvantaged populations throughout the corridor rather than separating different groups. Conversely, rivers with scenic waterfront views may attract advantaged populations to both banks, creating concentrations of similar groups on either side of the barrier. In both cases, urban barriers create similar conditions on both sides, and these attraction or repulsion effects may have a stronger influence on residential choice than any preference to use barriers as a way to separate from dissimilar groups.

Our methodological approach to generating synthetic urban partitions represents a deliberate trade-off between realism and statistical rigor that directly impacts our results. The way synthetic partitions are defined fundamentally affects the distribution of synthetic overlap scores against which we compare observed patterns. Creating more realistic or plausible synthetic partitions would require incorporating additional spatial rules and assumptions, which could also introduce potential bias through circular reasoning. Assumptions about plausible urban layouts could artificially produce the alignment (or misalignment) patterns we seek to detect. Therefore, we chose a plain approach that controls only for the statistical properties directly affecting our overlap metric—specifically, the number of urban fragments and their size distribution, since more evenly distributed fragment sizes naturally yield better overlap scores. By avoiding additional spatial assumptions, we essentially treat each city as a blank canvas and redraw urban frontiers randomly while preserving only the size characteristics of the actual fragmentation pattern.

Several limitations of our study should be acknowledged. First, our analysis focuses specifically on residential segregation of non-EU migrants, while income disparities may be more important for spatial sorting in many European contexts [26]. Second, despite efforts to build standardized data across countries, notable methodological differences remain that could influence our results [19]. These differences include variations in spatial resolution, where some countries like the Netherlands provide data natively at 100×100m resolution, while others, such as France, require upscaling from coarser administrative zones (approximately 300×300m) that were subsequently disaggregated into smaller grid cells. From a statistical perspective, larger original zones tend to have migrant shares closer to the city average due to their larger sample sizes, which reduces the extreme values that drive segregation patterns. Additionally, the operational definition of migration background varies significantly across countries: for instance, a person born in the Netherlands to parents from outside the EU who acquires Dutch citizenship retains their migration background classification, whereas in France, the same individual would no longer be classified as having a migration background upon acquiring French nationality as an adult. These differences in definition and resolution directly impact measurements of segregation intensity. Our alignment analysis is less sensitive to these variations since we focus on spatial layout patterns rather than absolute migrant shares, and the demographic regions we identify are typically orders of magnitude larger than even the coarsest spatial resolution in our dataset. Third, our data is from 2011, and segregation patterns may have changed in the subsequent years, particularly given the significant migration flows to Europe since the mid-2010s.

The underlying mechanisms driving the country-specific patterns we observe remain unclear and warrant further investigation in future research. In particular, understanding why Dutch and German cities exhibit stronger alignment between urban barriers and social frontiers compared to other European countries could provide valuable insights into how urban planning traditions and infrastructure development influence residential segregation patterns. To uncover these mechanisms, future research could investigate other spatial features that might overlap with segregation patterns in European cities. For instance, urban barriers might delineate housing development patterns, where plots of land bounded by railways, motorways, or waterways were developed with distinct housing types or at different time periods, potentially creating areas with varying affordability and accessibility that indirectly influence residential sorting patterns. The methodological framework developed in this study — comparing observed spatial overlap against synthetic null distributions — could be readily adapted to assess alignment between demographic patterns and partitions based on housing quality, concentration of social housing, building age, or access to amenities. Such approaches might reveal more consistent patterns that explain the spatial distribution of social groups in European cities, where segregation clearly exists but follows a different spatial logic than the barrier-bounded patterns observed in North American contexts.

4 Methods

Our study investigates whether residential segregation patterns align with urban fragmentation across European cities. We employ a Monte Carlo approach to test the statistical significance of the observed spatial overlap between demographic regions and urban fragments defined by urban barriers. For each of the 520 cities in our sample, we generate 200 synthetic fragmentation patterns and compare the actual spatial alignment against this null distribution. This methodology allows us to determine whether urban barriers act as social frontiers more than would be expected by chance.

4.1 Data Sources and Preparation

Our analysis requires two datasets: demographic data to identify patterns of residential segregation and geographic data to define urban barriers. The following subsections detail our data sources and the preparation steps we undertook to create standardized, comparable spatial partitions of residential segregation and urban fragmentation across all 520 cities in our sample.

4.1.1 Demographic dataset

This study used a dataset mapping the spatial distribution of migrants across urban areas in eight European countries (France, Germany, Ireland, Italy, the Netherlands, Portugal, Spain, and the UK) derived from harmonized 2011 Census data [19]. The dataset was created by processing ad hoc extractions from National Statistical Institutes, organizing population data by citizenship and/or country of birth into a uniform spatial grid with cells of 100×100 meters. The population was classified into two categories: migrants from outside the EU and the rest (including non-migrants and EU migrants). The definition of migrants varied across countries, based on citizenship criteria in Italy and France, and on country of birth in the remaining countries (Germany, Ireland, the Netherlands, Portugal, Spain, and the UK).

4.1.2 Study area

We define the spatial extent of urban areas through population density analysis. First, we calculate the average population density within a 400-meter radius of each cell, applying a smoothing function that weighs nearby populations inversely to the square of the distance. The 400-meter radius was selected based on a qualitative assessment of different smoothing parameters to create cohesive urban areas while preserving meaningful local variation in density. Cells exceeding a density threshold of 1,000 people per km^2 are classified as high-density areas. This conservative threshold accounts for edge effects, which cause the moving average calculation to yield densities approximately half those of central areas in peripheral cells. Adjacent high-density cells are then aggregated into contiguous urban areas, and only areas with a total population exceeding 50,000 inhabitants are retained, consistent with the established definition of urban cores in the literature [27]. We define the study area using a convex hull around each urban area to ensure we capture any urban barriers that might exist at the periphery of settlements. This methodological approach ensures our analysis focuses on densely populated urban environments while maintaining consistency across the 520 cities in our sample.

4.1.3 Regionalization method for constructing demographic regions

We identify homogeneous demographic regions using the regionalization method developed in [20]. This method applies a spatial moving average to migrant proportion data to filter out small-scale variations. It then uses spatially-constrained agglomerative clustering, which ensures only adjacent cells merge into regions, maintaining spatial contiguity throughout. The result is a set of contiguous demographic regions for each city, with each region representing an area of relatively homogeneous migrant population characteristics, enabling direct comparison with urban fragments defined by urban barriers.

4.1.4 Identification of urban barriers and creation of urban fragments

We extract urban barriers from the OpenStreetMap database, which include railways, motorways, and waterways [28]. We then systematically partition the demographic grid: each contiguous group of cells that is fully separated from other groups by one or more urban barriers constitutes an urban fragment. This fragmentation process creates a spatial partition that we then compare against the independently derived demographic regions to assess potential alignment between urban barriers and social frontiers.

4.2 Quantifying alignment between social frontiers and urban barriers

To assess whether residential segregation patterns align with urban fragmentation beyond what would be expected by chance, we employ a two-step approach. First, we develop a purity score metric to quantify the spatial overlap between demographic regions and urban fragments within each city. Second, we establish statistical significance by comparing observed alignment scores against a null distribution generated through Monte Carlo simulations using synthetic urban partitions.

4.2.1 Measuring the alignment between two spatial partitions using the purity score

We measure the alignment between residential segregation patterns and urban fragmentation using a purity score metric η (equation 1). For each urban fragment m (constructed in subsection 4.1.4) in the set M of all urban fragments, we identify the demographic region l (constructed in subsection 4.1.3) in the set of all demographic regions L that contains the largest share of the fragment's population. The overall purity score sums these maximum population overlaps across all fragments M, normalized by the total population pop_{city}.

$$\eta_m = \frac{1}{pop_{city}} \sum_{m \in M} \max_{l \in L}(pop(m \cap l)) \tag{1}$$

The purity score can be interpreted as a measure indicating how well one can reconstruct demographic regions using urban fragments as building blocks. It ranges from 0 to 1, with 1 implying that demographic regions can be perfectly reconstructed from urban fragments. This metric is calculated for each city as well as for each of the synthetic partitions.

4.2.2 Generation of synthetic partitions

We generate synthetic urban partitions by randomly placing seed points within the city area, matching the number of fragments observed in the actual urban structure. We then apply Voronoi tessellation, assigning each population grid cell to its nearest seed to form distinct urban fragments. We use transformed Gaussian random fields to draw these random points, providing flexible control over spatial patterns. A Gaussian random field $Z(x, y)$ assigns random values to spatial locations (x, y), where any finite collection of field values follows a multivariate normal distribution. The field is characterized by its covariance function:

$$\text{Cov}[Z(x_k, y_k), Z(x_{k'}, y_{k'})] = \sigma^2 \exp\left(-\frac{h^2}{\lambda^2}\right) \tag{2}$$

$$P(x) = \frac{[Z(x, y) + |\min(Z)|]^\zeta}{\sum_k [Z(x_k, y_k) + |\min(Z)|]^\zeta} \tag{3}$$

where h is the distance between points, λ is the length scale parameter controlling spatial correlation, and ζ is the power exponent. Equation 2 defines how field values $Z(x, y)$ at different locations covary—larger λ values produce smoother, more clustered patterns while smaller λ values create more dispersed arrangements. The power transformation in equation 3 converts the field into probabilities for seed placement, where higher ζ values increase clustering around field maxima.

To ensure that synthetic structures accurately match the size distribution of actual urban fragments, we implement a two-stage process as follows.

First, we generate 500 fields with varying λ and ζ and construct corresponding synthetic partitions, optimizing the length scale controlling spatial autocorrelation (equation 2) and the power exponent regulating point concentration (equation 3). Their optimal values are determined by minimizing the Wasserstein distance between synthetic and observed fragment size distributions (equation 4). We measure the Wasserstein distance W_1 by ranking fragments by descending size, constructing the cumulative distribution function of fragment size in the actual case A and the synthetic case S, and summing the difference between the two over every rank r.

$$W_1(A, S) = \sum_{r \in R} |CDF_A(r) - CDF_S(r)| \tag{4}$$

Second, once optimal parameters are identified, we generate 500 synthetic partitions using these optimized parameters and select the top 200 partitions with the smallest Wasserstein distances to the actual fragmentation partition. This approach creates multiple statistically equivalent alternatives that preserve the size distribution of the actual urban fragmentation partition.

4.2.3 Statistical significance of alignment between residential segregation and urban fragmentation

We test the statistical significance of the alignment between the demographic partition and the fragmentation partition by comparing the observed purity score against the distribution of scores from 200 synthetic fragmentation partitions. For each city, we calculate the quantile rank of the actual purity score within the distribution of synthetic purity scores, which provides a standardized measure of statistical significance at the city level.

The quantile rank q for city j is computed as:

$$q_j = \frac{|s \in S_j : \eta_s < \eta_{observed,j}|}{|S_j|} \tag{5}$$

where $\pi_{observed,c}$ represents the observed purity score for city j, S_j denotes the set of 200 synthetic purity scores for that city, and $|\cdot|$ indicates set cardinality. This quantile rank ranges from 0 to 1, with values approaching 1 indicating that the observed alignment is exceptionally high compared to what would be expected under random spatial arrangements of urban fragments.

A quantile rank above 0.95 suggests that the observed alignment between residential segregation and urban fragmentation occurs in fewer than 5 % of random scenarios, indicating statistically significant alignment at the $\alpha = 0.05$ level. Conversely, quantile ranks below 0.05 indicate significantly less alignment than expected by chance, suggesting that urban barriers actively disagree with social frontiers.

This quantile-based approach provides several methodological advantages: it naturally accounts for the varying number and size distribution of urban fragments across cities, it makes no distributional assumptions about the purity scores, and it provides an intuitive interpretation where the quantile rank directly corresponds to the probability that a randomly generated urban fragmentation would yield lower alignment than the observed pattern.

4.3 Post-hoc statistical testing

After conducting our Monte Carlo analysis to assess alignment between residential segregation and urban fragmentation in individual cities, we perform two additional statistical procedures to analyze patterns across our entire sample. First, we test whether the distribution of quantile ranks across all cities deviates significantly from what would be expected under the null hypothesis of no systematic relationship (subsection 4.3.1). Second, we examine country-specific effects on alignment patterns while controlling for city-level characteristics that might influence this relationship (subsection 4.3.2).

4.3.1 Benjamini-Yekutieli test

In section 2.1, we examine which bins deviate from what we would expect under the null hypothesis of uniformity. A key challenge in this analysis is the dependence between bins: if one bin has significantly more observations than expected, other bins have, by construction, fewer observations. The Benjamini-Yekutieli procedure handles such dependencies between tests [21]. This procedure controls the false discovery rate (FDR) — the expected proportion of false positives among all rejected null hypotheses — when performing multiple hypothesis tests, which we set to FDR =0.05 in this study. This approach provides a statistical framework that accounts for the inherent dependency in our quantile rank distribution analysis. The implementation involves five steps:

1. Conducting Chi-Square tests comparing observed versus expected counts for each bin b, and computing the p-values per bin p_b.
2. Ordering the resulting p-values from smallest to largest: $p^{(1)} \leq p^{(2)} \leq \ldots \leq p^{(20)}$, where κ denotes the rank
3. Calculating Benjamini-Yekutieli critical values τ_κ (see equation 6)
4. Finding the largest $\hat{\kappa}$ where $p^{(\hat{\kappa})} \leq \tau_{\hat{\kappa}}$
5. Rejecting the null hypothesis for all bins with p-values $\leq p^{(\hat{\kappa})}$

$$\tau_\kappa = \frac{\kappa}{20} \times \frac{FDR}{c} \quad \text{where} \quad c = \sum_{\kappa'=1}^{20} \frac{1}{\kappa'} \tag{6}$$

4.3.2 Beta regression

To measure the absolute effect of each country on the alignment between urban barriers and social frontiers, we employ beta regression analysis while controlling for city-level characteristics that might influence this relationship. Beta regression is particularly appropriate for our analysis because it is specifically designed for continuous response variables bounded between 0 and 1, exactly matching the properties of our quantile ranks [22]. This regression framework allows us to test whether country-specific effects remain statistically significant after accounting for segregation levels and decentralization patterns. Equation 8 provides the regression specifications, where Y_{jc} represents the response variable (quantile rank) for city j in country c. The variable follows a beta distribution with mean μ_j and precision parameter ϕ. We use logit as the link function $g(\cdot)$ connecting the linear predictor $\mathbf{x}_j^T\boldsymbol{\beta}$ to the mean μ_j, which ensures predicted values remain within the (0,1) interval while maintaining linearity on the transformed scale.

$$y_{jc} \sim \text{Beta}(\mu_{jc}, \phi) \tag{7}$$

$$\text{logit}(\mu_{jc}) = \alpha_c + \beta_D \cdot D_{jc} + \beta_\delta \cdot \delta_{jc} + \varepsilon_{jc} \tag{8}$$

We control for segregation intensity and decentralization, denoted respectively by D and δ in equation 8. Both control variables are centered around their respective means to enable meaningful coefficient interpretation, since baseline scenarios with zero decentralization or segregation are neither realistic nor meaningful in urban contexts.

We measure the intensity of residential segregation in an urban area using the dissimilarity index [29]. This index noted D is measured using equation 9 where u_k and v_k represent respectively the migrant and the non-migrant population in unit k, while the total migrant and non-migrant population in the urban area are noted U and V respectively. This index ranges from 0 (no segregation) to 1 (complete segregation), providing a standardized measure of spatial separation. We use this established measure to maintain consistency with prior studies employing the same dataset [23, 30, 31].

$$D = \frac{1}{2}\sum_k \left| \frac{u_k}{U} - \frac{v_k}{V} \right| \tag{9}$$

We also control for decentralization, which captures the fundamental spatial organization of urban development. Cities can be organized along a spectrum from monocentric structures—where development radiates outward from a single dominant center—to multicentric structures—where development is distributed across multiple distinct centers or nodes. The decentralization is computed using the average distance from any cell k located at position (x_k, y_k) to the city's center of gravity (x_g, y_g) weighted by the cell's population pop_k (see equation 10). The city's center of gravity is also population-weighted. This indicator is linearly correlated to the city size. To enable cross-city comparisons, we normalize this measure by dividing it by the theoretical average distance from the city center ADC_{circle} of a perfectly circular city with the same area S (equation 11). The normalization in equation 12 yields the decentralization indicator δ, a dimensionless metric that allows for meaningful comparison of decentralization patterns across urban areas of different sizes and shapes.

$$ADC = \frac{\sum_k \sqrt{(x_k - x_g)^2 + (y_k - y_g)^2} \cdot pop_k}{\sum_k pop_k} \tag{10}$$

$$ADC_{circle} = \frac{2}{3} \cdot \sqrt{\frac{S}{\pi}} \tag{11}$$

$$\delta = \frac{ADC}{ADC_{circle}} \tag{12}$$

The country fixed effects are denoted by α_c. Our specification excludes an intercept term and retains all country fixed effects, allowing us to assess whether each country exhibits positive or negative effects on the alignment phenomenon rather than comparing countries against a single reference.

Acknowledgments. This work is supported by the TU Delft AI initiative. Research interns Thibault Joumard, Martin de Jaeghere, and Romuald Winandy-Proust assisted in the preprocessing stage of the analysis. We thank the National Statistical Institutes ISTAT (Italy), INE (Spain), DESTATIS (Germany), INSEE (France), CBS (Netherlands), as well as the European Commission's Joint Research Centre (JRC) for providing the original data from the 2011 Census.

References

[1] Reardon, S.F., O'Sullivan, D.: Measures of spatial segregation. Sociological methodology **34**(1), 121–162 (2004)

[2] Greft, S., Fortuijn, J.D.: Multiple disadvantage of older migrants and native Dutch older adults in deprived neighbourhoods in Amsterdam, the Netherlands: a life course perspective. GeoJournal **82**, 415–432 (2017) https://doi.org/10.1007/s10708-015-9691-x

[3] Owens, A.: Income segregation between school districts and inequality in students' achievement. Sociology of Education **91**, 1–27 (2018) https://doi.org/10.1177/0038040717741180

[4] Nieuwenhuis, J., Tammaru, T., Ham, M., Hedman, L., Manley, D.: Does segregation reduce socio-spatial mobility? Evidence from four European countries with different inequality and segregation contexts. Urban Studies **57**, 176–197 (2020) https://doi.org/10.1177/0042098018807628

[5] Tóth, G., Wachs, J., Di Clemente, R., Jakobi, Á., Ságvári, B., Kertész, J., Lengyel, B.: Inequality is rising where social network segregation interacts with urban topology. Nature communications **12**(1), 1143 (2021)

[6] Ananat, E.O.: The wrong side (s) of the tracks: The causal effects of racial segregation on urban poverty and inequality. American Economic Journal: Applied Economics **3**(2), 34–66 (2011)

[7] Mahajan, A.: Highways and segregation. Journal of Urban Economics **141**, 103574 (2024) https://doi.org/10.1016/j.jue.2023.103574

[8] Noonan, D.S.: Neighbours, barriers and urban environments: Are things 'different on the other side of the tracks'? Urban Studies **42**(10), 1817–1835 (2005) https://doi.org/10.1080/00420980500231720 https://doi.org/10.1080/00420980500231720

[9] Kramer, R.: Defensible spaces in Philadelphia: Exploring neighborhood boundaries through spatial analysis. RSF: The Russell Sage Foundation Journal of the Social Sciences **3**(2), 81–101 (2017)

[10] Kramer, R.: Testing the role of barriers in shaping segregation profiles: The importance of visualizing the local neighborhood. Environment and Planning B: Urban Analytics and City Science **45**(6), 1106–1121 (2018)

[11] Roberto, E., Korver-Glenn, E.: The spatial structure and local experience of residential segregation. Spatial Demography **9**(3), 277–307 (2021)

[12] Ballard, R., Hamann, C.: Urban Socio-Economic Segregation and Income Inequality: A Global Perspective, pp. 91–109. Springer, Cham (2021). Chap. 5. Income inequality and socio-economic segregation in the city of Johannesburg

[13] Chyn, E., Haggag, K., Stuart, B.A.: The effects of racial segregation on intergenerational mobility: Evidence from historical railroad placement. Technical report, National Bureau of Economic Research (2022)

[14] Egede, L.E., Walker, R.J., Campbell, J.A., Linde, S., Hawks, L.C., Burgess, K.M.: Modern day consequences of historic redlining: finding a path forward. Journal of general internal medicine **38**(6), 1534–1537 (2023)

[15] Joshi, A., Horn, B.P., Berrens, R.P.: Contemporary differences in residential housing values along historic redlining boundaries. Real Estate Economics **52**(2), 514–551 (2024)

[16] Maharaj, B.: South African Urban Planning in the Twentieth and Twenty-First Centuries—Continuities between the Apartheid and Democratic eras? Urban and regional planning and development: 20th century forms and 21st century transformations, pp. 101–112. Springer, Cham (2020)

[17] Schelling, T.C.: Dynamic models of segregation. Journal of mathematical sociology **1**(2), 143–186 (1971)

[18] Roberto, E.: The spatial proximity and connectivity method for measuring and analyzing residential segregation. Sociological Methodology **48**(1), 182–224 (2018)

[19] Alessandrini, A., Natale, F., Sermi, F., Vespe, M.: High resolution map of migrants in the EU. Technical report, Joint Research Centre, Luxembourg (Luxembourg) (2017)

[20] Spierenburg, L., van Cranenburgh, S., Cats, O.: Evolution of residential segregation patterns in the Netherlands between 2015 and 2020. Cities **150**, 105038 (2024) https://doi.org/10.1016/j.cities.2024.105038

[21] Benjamini, Y., Yekutieli, D.: The control of the false discovery rate in multiple testing under dependency. Annals of statistics, 1165–1188 (2001)

[22] Ferrari, S., Cribari-Neto, F.: Beta regression for modelling rates and proportions. Journal of applied statistics **31**(7), 799–815 (2004)

[23] Benassi, F., Bonifazi, C., Heins, F., Lipizzi, F., Strozza, S.: Comparing residential segregation of migrant populations in selected European urban and metropolitan areas. Spatial Demography **8**, 269–290 (2020)

[24] Grannis, R.: Chapter 5: selection and Influence. From the Ground Up: Translating Geography into Community through Neighbor Networks, pp. 48–58. Princeton University Press, Princeton (2009)

[25] Roberto, E., Hwang, J.: Barriers to integration: Institutionalized boundaries and the spatial structure of residential segregation. In: Annual Meeting of the Population Association of America, Chicago, Illinois (2017)

[26] Musterd, S., Marcińczak, S., van Ham, M., Tammaru, T.: Socioeconomic segregation in European capital cities. increasing separation between poor and rich. Urban Geography **38**(7), 1062–1083 (2017) https://doi.org/10.1080/02723638.2016.1228371

[27] Dijkstra, L., Poelman, H., Veneri, P.: The EU-OECD definition of a functional urban area. OECD Regional Development Working Papers **No.2019** (2019)

[28] OpenStreetMap: OpenStreetMap data. Last accessed on May 5, 2025 from https://www.openstreetmap.org/ (2025)

[29] Duncan, O.D., Duncan, B.: A methodological analysis of segregation indexes. American sociological review **20**(2), 210–217 (1955)

[30] Benassi, F., Naccarato, A., Iglesias-Pascual, R., Salvati, L., Strozza, S.: Measuring residential segregation in multi-ethnic and unequal european cities. International Migration **61**(2), 341–361 (2023)

[31] Marcińczak, S., Mooses, V., Strömgren, M., Tammaru, T.: A comparative study of immigrant-native segregation at multiple spatial scales in urban Europe. Journal of Ethnic and Migration Studies **49**(1), 43–65 (2023)